Methuen's Monographs on Physical Subjects

General Editor: B. L. Worsnop, B.Sc., Ph.D.

WIRELESS RECEIVERS

WIRELESS RECEIVERS

THE PRINCIPLES OF THEIR DESIGN

BY

C. W. OATLEY, M.A., M.Sc.

KING'S COLLEGE, LONDON

WITH A PREFACE BY

O. W. RICHARDSON, F.R.S.

YARROW RESEARCH PROFESSOR OF THE ROYAL SOCIETY
NOBEL LAUREATE IN PHYSICS, 1928

WITH 41 DIAGRAMS

METHUEN & CO. LTD.
36 ESSEX STREET W.C.
LONDON

FOURTH EDITION

First Published	. .	*May 12th 1932*
Second Edition	. .	*January 1940*
Third Edition	. .	*October 1942*
Fourth Edition	. .	*1945*

PRINTED IN GREAT BRITAIN

GENERAL PREFACE

THIS series of small monographs is one which should commend itself to a wide field of readers.

The reader will find in these volumes an up-to-date résumé of the developments in the subject considered. The references to the standard works and to recent papers will enable him to pursue further those subjects which he finds of especial interest. The monographs should therefore be of great service to physics students who have examinations to consider, to those who are engaged in research in other branches of physics and allied sciences, and to the large number of science masters and others interested in the development of physical science who are no longer in close contact with recent work.

From a consideration of the list of authors it is clear that the reader need have no doubt of the accuracy of the general accounts found in these volumes.

O. W. RICHARDSON

KING'S COLLEGE
LONDON

PREFACE

AN attempt has been made to present in the following pages a fairly detailed account of the fundamental principles involved in the design of a wireless receiver. Most of the material has been gathered from original papers, detailed references to which are given in the bibliography at the end of the book. Limitations of space have made it necessary to omit many important aspects of the subject, but it seemed better to treat selected portions thoroughly, rather than to give a vague outline of the whole. Perhaps the most serious omission is that of any mention of the Superheterodyne Receiver. However, a few references to this subject have been added at the end of the bibliography.

I wish to express my warmest thanks to Professor E. V. Appleton, both for his encouragement and for his kindness in reading through the manuscript and offering helpful criticisms while the book was passing through the press.

C. W. O.

WHEATSTONE LABORATORY
KING'S COLLEGE
LONDON, *March, 1932*

CONTENTS

WIRELESS RECEIVERS

CHAPTER I

INTRODUCTORY

WHEN wireless communication is established between two stations, radiations sent out by one station travel through the æther and act upon the aerial-earth system of the other station, where they cause an alternating electric potential to be set up between two points of this system. The wave-form of the alternating potential will correspond to that of the æther waves and, in the simplest possible case, both will be sinusoidal. However, unless the wave-form vary with time, it will clearly not be possible to convey any message from one station to the other. One possible method of variation, which is used in wireless telegraphy, is to divide the originally continuous wave into a series of discrete wave trains, which, according to their length, may represent dots or dashes. The message can then be conveyed by means of the Morse Code. In wireless telephony, on the other hand, the continuity of the wave is retained, but the amplitude is made to vary with time in such a way that the wave-form of the variations is a reproduction of the acoustic wave-form of the sounds which compose the message. Such a wave is said to be modulated.

Let it be assumed that the emitting station is radiating

at a wave-length λ, and that the signal which is being transmitted is a sustained note of acoustic frequency f.

Then the wave-form of the alternating potential V, which is set up in the aerial-earth system of the receiver, may be represented diagrammatically as in Fig. 1, where potential V is plotted against time t. This represents a sinusoidal electromotive force (E.M.F.) the amplitude of which varies from $A - m$ to $A + m$, the variations also taking place sinusoidally with respect to time. The frequency f of the amplitude variations will be the same as that of the signal note which is being transmitted,

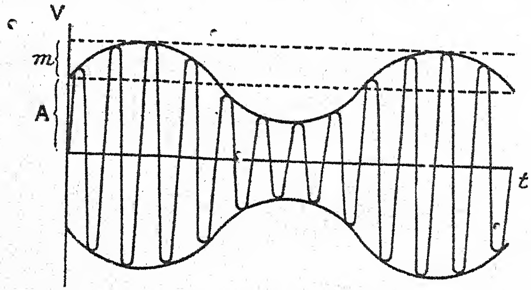

Fig. 1.

while the frequency F of the sinsuoidal E.M.F. will be given by

$$F\lambda = U \quad . \quad . \quad . \quad . \quad (1)$$

where U is the velocity of wireless waves in free space. If λ be measured in metres and F in kilocycles per second, then

$$F = 3 \times 10^5/\lambda \quad . \quad . \quad . \quad (2)$$

The alternating E.M.F. which is set up in the aerial-earth system will cause corresponding variations of potential in other stages of the receiver and, as a result, the loud speaker will emit an audible note of frequency f. In the following pages the various stages of a receiver will be considered in detail, and particular attention will be paid

to the principles of design which must be observed if the note emitted by the loud speaker is to be a faithful copy of the note which falls upon the microphone at the emitting station.

A wave of the form shown in Fig. 1 is known as a *modulated* sine wave; f is known as the *modulation frequency* and 100 m/A as the *percentage of modulation.* Then if

$$\omega = 2\pi F \text{ and } p = 2\pi f^* \quad . \quad . \quad (3)$$

the wave-form of Fig. 1 may be expressed mathematically by the equation

$$V = (A + m \sin pt) \sin \omega t.$$

This may be re-written as

$$\begin{aligned} V &= A \sin \omega t + m \sin pt \sin \omega t \\ &= A \sin \omega t + \frac{m}{2} \cos(\omega - p)t - \frac{m}{2} \cos(\omega + p)t \quad . \quad (4) \end{aligned}$$

It thus appears that the E.M.F. set up in the aerial-earth system consists of three distinct components, which we may suppose to have been produced by the simultaneous action on the aerial of three distinct sinusoidal electromagnetic waves. The wave, the pulsatance of which is ω is known as the *Carrier Wave*, while the waves having pulsatances $\omega + p$ and $\omega - p$ respectively, are known as the *Side Bands*. In wireless telephony, the modulation frequency f will lie within the audible range, and will therefore not exceed about 10,000 cycles per second. Frequencies which do not exceed this limit are referred to as *audio* or *low* frequencies. On the other hand, the frequency F of the carrier wave will not usually be less than 100 kilocycles per second. Such frequencies are referred to as *radio* or *high* frequencies. Since then f is always small compared with F, it follows that the frequencies of the side bands will lie in the radio frequency

* The frequency of an oscillation multiplied by 2π is known as the *pulsatance* of the oscillation.

range. The abbreviations H.F. and L.F. are commonly employed to denote high frequency and low frequency respectively.

Function of a Wireless Receiver

The function of the wireless receiver proper is to convert the E.M.F. which is generated in the aerial-earth system, and which is of the form represented by equation (4), into a sinusoidal E.M.F. of frequency f, which can be applied to the loud speaker. Furthermore, the receiver must be capable of furnishing sufficient electrical power to operate the loud speaker at the required strength. In general, the receiver consists of four distinct parts or "stages." The E.M.F. from the aerial is applied to a *High-Frequency Amplifier*, the function of which is to increase the magnitude without changing the form of this E.M.F. The amplified E.M.F. is next applied to a *Detector Stage*. Whereas the input E.M.F. to this stage consists of three radio frequency components, the output consists of an E.M.F. of frequency f. Next follows a *Low-Frequency Amplifier*, the function of which is to increase the magnitude without changing the form of the low frequency which is applied to it. It will be noticed that the functions of the above three stages have been described in terms of E.M.F., while nothing has been said about the quantity of current flowing. This is because each stage absorbs an extremely small amount of power from the preceding one, so that the currents flowing are very small, although the E.M.F.s may be quite large. As a matter of fact, none of the above stages would be capable of furnishing very much power to the one following it. Hence the necessity for the last stage of the receiver, which is known as the *Power Stage*. The output from this stage is controlled by the E.M.F. applied to it by the low-frequency amplifier, but very little power is absorbed from the latter. The function of the power stage is to deliver sufficient power to operate the loud speaker.

From the above description, it will be clear that neither the H.F. amplifier nor the L.F. amplifier is essential to a receiver, and either or both may be omitted if sufficient volume of sound can be obtained without them..

Types of Distortion

Before discussing the principles of design which should be followed if the receiver is to cause no distortion of the signal, we must consider briefly the types of distortion which may occur. Since this book is concerned only with wireless receivers, it will be assumed that both the transmitting station and the loud speaker are perfect, and so do not distort the signal.

Consider first the case where the signal to be received is a pure, sustained audio-frequency note. Then, owing to imperfections in the receiver, it may happen that the amplitude of the note emitted by the loud speaker is not proportional to the amplitude of the note sounded in front of the microphone. This type of distortion is known as *Amplitude Distortion*, and when it is present, harmonics of the original note are introduced so that the note emitted by the loud speaker is no longer pure.

When amplitude distortion is absent, the amplitudes of these two notes will be proportional, but it may still happen that the constant of proportionality is different for different frequencies. In this case, *Frequency Distortion* is said to occur. This type is not so objectionable as amplitude distortion for two reasons. In the first place, frequency distortion occurring at one stage of a receiver can often be corrected at a later stage. Secondly, the human ear is very insensitive to change of volume of sound, and so is unable to detect slight frequency distortion. Both types of distortion mentioned above may occur simultaneously.

So far, it has been assumed that the signal is a sustained sinusoidal note. In practice, of course, this will rarely, if ever, be the case. Let us consider next what will happen if the signal consist of a sustained note, the

wave-form of which is not sinusoidal. Such a note may, by Fourier analysis, be split up into a series of sinusoidal components. Each of these components will cause the transmitter to radiate a definite pair of side-bands of corresponding frequencies, while the carrier wave will remain as before, since it is not affected by the microphone output. Thus, as far as the transmitter is concerned, we may treat each component of the complex note as a separate entity, which is not affected by the presence of the other components. When the various parts of a receiver have been considered in detail, it will be seen that this statement also holds true for a receiver, provided neither frequency nor amplitude distortion be present. There is still, however, the possibility that the phases of the various components will be changed with respect to each other. We may refer to this as *Phase Distortion.* In a recent paper (1), van der Pol has described an ingenious arrangement whereby it is possible to change the phases of the components of a complex E.M.F. without altering their relative amplitudes. By connecting this device in a loud speaker circuit, he was able to show that the human ear is unable to detect phase distortion, even though the change in phase of one component relatively to the others approaches 360°. Indirect evidence had previously led Helmholtz to a similar conclusion.

From what has been said, it will be clear that, so long as we are concerned with distortion which may arise when the microphone output is a sustained alternating E.M.F., it will be sufficient if we consider amplitude and frequency distortion in the simple case when the microphone output is a sinusoidal E.M.F. the frequency of which may lie anywhere within the audible range. Accordingly, this simple case is the one which will be treated in the following chapters.

There remains for consideration the case when the microphone output E.M.F. is not sustained or periodic, but is in the form of an irregular pulse, which we may term a *Transient* E.M.F. Any distortion of the form of

such an E.M.F. may be termed *Transient Distortion.* A great deal of work remains to be done on this subject, but two facts have been fairly well established ; firstly, that the transient distortion in a well-designed receiver is probably small compared with that introduced by the best of present-day loud speakers ; and secondly, that transient distortion in a good receiver does not prevent the quality of reproduction of speech and music from being extremely good. For these reasons transient distortion will not be considered in the present book.

CHAPTER II

THE TRIODE AND ITS EQUIVALENT CIRCUIT

Characteristics of a Triode

IN a companion volume of the present series (2), a detailed account has been given of the construction and internal action of thermionic valves. Although such valves are of fundamental importance in wireless receivers, we shall here be concerned only with their properties when connected in electric circuits, and shall not enquire into the reasons for these properties. From this point of view, the behaviour of any particular valve is best summarised by a series of curves showing the relation between the currents flowing to or from the various electrodes and the potentials applied to these electrodes. Such curves are termed the *Characteristics* of the valve, and, by convention, the potentials of all the electrodes are measured with reference to the potential of the negative end of the filament, which is taken to be zero. In "Indirectly Heated" valves the kathode is taken to be at zero potential.

The present chapter will be devoted to a consideration of the properties of three-electrode valves or triodes.

Since the anode current of a triode is a function of both anode and grid potentials, the complete current-potential characteristic can only be represented by a surface in three dimensions. If, however, one of the parameters be kept constant, a plane curve can be drawn showing the relation between the other two. Two sets of such curves are shown in Fig. 2 (*a*) and (*b*).

respectively. In Fig. 2 (a) curves are drawn to indicate the relation between anode current (I_a) and grid potential (V_g), for various constant values of anode potential (V_a). Similarly, Fig. 2 (b) indicates the relation between anode current and anode potential, when grid potential is kept constant.

Suppose a triode to be operated under conditions represented by point A in Fig. 2 (a). Now let the anode potential be increased by an amount δV_a, while the grid potential is decreased simultaneously by an amount

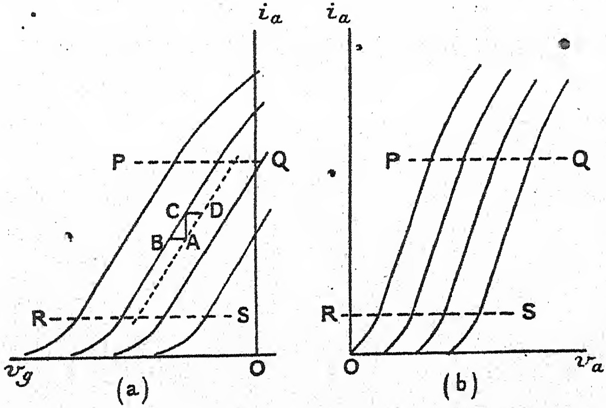

FIG. 2.

δV_g, so that the point of operation moves to B and the anode current remains unchanged. Then the limit of the fraction $\frac{\delta V_a}{\delta V_g}$ as δV_a and δV_g become indefinitely small, is known as the *Amplification Factor* of the triode, and is usually denoted by the symbol μ.

Next suppose that, beginning at point A, the anode potential increases by an amount δV_a, while the grid potential remains constant, so that the point of operation moves to C and the anode current increases by an amount

2

δI_a. Then the limit of $\frac{\delta V_a}{\delta I_a}$, for indefinitely small increments, is denoted by the symbol ρ and is known as the *Impedance, Differential Resistance* or *Slope Resistance* of the triode. Of these terms, the first is a misnomer, since it suggests the presence of a reactive component, whereas the quantity concerned is of the nature of a pure resistance. The second term is mathematically correct, but is somewhat cumbersome, and for these reasons the term Slope Resistance will be used in the present book. This quantity must be carefully distinguished from the *Direct Current Resistance* of the triode. This latter quantity, for any point of operation, is defined to be the ratio of the *total* anode potential to the *total* anode current, and clearly, this is, in general, a very different thing from the ratio of a small increment of anode potential to the corresponding increment of anode current.

Finally, commencing from point A, suppose the grid potential to increase by an amount δV_g, while the anode potential remains constant, so that the point of operation moves along the curve to D and the anode current increases by an amount δI_a. Then the limit of the ratio $\frac{\delta I_a}{\delta V_g}$ for indefinitely small increments, is known as the *Mutual Conductance* of the triode, and will be denoted by the symbol κ. This term is accepted generally and will be retained here, though it is somewhat misleading, since it is not connected with any "Mutual" or "Reciprocal" action.

Putting the above definitions into mathematical form, we may write,

$$\mu = \frac{\partial V_a}{\partial V_g}, \qquad \rho = \frac{\partial V_a}{\partial I_a}, \qquad \kappa = \frac{\partial I_a}{\partial V_g},$$

whence it is obvious that

$$\kappa = \frac{\mu}{\rho}.$$

Since the above definitions have been given in terms of infinitesimal variations, they can be applied to any point on a triode characteristic surface. The importance of the quantities defined, however, lies in the fact that, over a considerable portion of the working characteristic of the triode, they are approximately constant. For example, this will be the case for the curves illustrated in Fig. 2 (*a*) and (*b*) over the regions lying between the dotted lines PQ and RS.

Fundamental Equation for a Triode

Consider next how variations of anode current depend upon variations of grid and anode potentials. With the above notation, we may write

$$\begin{aligned}\delta I_a &= \frac{\partial I_a}{\partial V_a}\delta V_a + \frac{\partial I_a}{\partial V_g}\delta V_g \\ &= \frac{1}{\rho}\delta V_a + \kappa\delta V_g \\ &= \frac{1}{\rho}\left(\delta V_a + \mu\delta V_g\right) \quad . \quad . \quad . \quad (5)\end{aligned}$$

Provided the conditions under which the triode is used be such that the point representing the instantaneous anode potential, grid potential and anode current never strays from the straight portions of the characteristic curves, the amplification factor and slope resistance will remain constant and equation (5) can be re-written as

$$i_a = \frac{1}{\rho}\left(v_a + \mu v_g\right) \quad . \quad . \quad . \quad (6)$$

since there is no longer any need to restrict the variations to small values. In this equation it is to be understood that the symbols i_a, v_a, v_g refer to variations of the quantities concerned from initial steady values.

The required initial steady values are obtained by placing suitable batteries or other equivalent devices in

the anode and grid circuits, and it will be assumed in what follows that, unless the contrary is specifically stated, these batteries are adjusted so that only the sensibly straight portions of the valve characteristics are used. In the foregoing account no mention has been made of the flow of current to the grid of a triode because, with one important exception to be considered later, the grid is always maintained at a negative potential and the grid current is then practically zero.

Equivalent Circuit of a Triode

Consider next the circuit shown in Fig. 3 (a), where an impedance Z and a battery of steady potential V

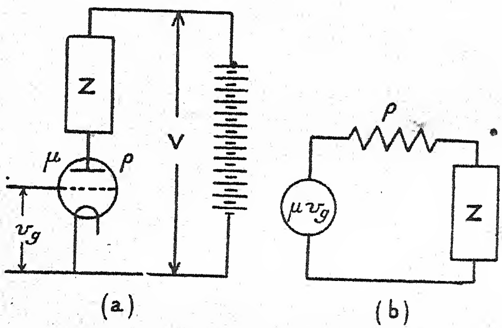

Fig. 3.

complete the anode circuit of a triode. Suppose the grid potential to be increased by an amount v_g, and let the resultant increase of current in the anode circuit be i_a. Owing to this increase in the current through Z, the potential of the an. de will change by an amonnt v_a. Furthermore, if v_z be the increase of potential difference across Z, we have, since V remains constant, $v_a = -v_z$. Then from (6)

$$\rho i_a + v_z = \mu v_g \quad . \quad . \quad . \quad (7)$$

Equation (7) shows that, for purposes of anode circuit calculations of potential and current changes, we may replace the triode by the equivalent circuit shown in Fig. 3 (*b*).

In future, unless the contrary is specifically stated, all varying currents and potentials will be assumed to be sinusoidal in form. Then impedances such as Z may be replaced by their equivalent series resistance R and series reactance X. Thus equation (7) may be written

$$\mu v_g = i_a(\rho + R + jX) \quad . \quad . \quad (8)$$

where j is the ordinary operator of alternating current theory. The numerical magnitude of Z will be denoted by $|Z|$ and is given by the equation

$$|Z| = \sqrt{R^2 + X^2}.$$

Effect of Inter-electrode Capacities

So far it has been assumed that a triode behaves as a pure resistance and possesses no reactive component. Actually this is not the case, since, between each pair of electrodes, there exists an electrical capacity, which, although small, is nevertheless very important at radio frequencies and has an appreciable effect at the higher audio frequencies. Taking these capacities into account and dealing only with *changes* of current and potential, a triode may be represented as in Fig. 4 (*a*) where Z_a and Z_g are the external impedances in the anode and grid circuits respectively and C_1, C_2, and C_3 are the inter-electrode capacities of the valve. A further complication arises from the fact that these small inter-electrode condensers are subject to considerable dielectric loss, and consequently an effective resistance should be included in series with each one. The theoretical treatment of this circuit has been developed by several writers, among whom may be mentioned Nichols (3), Miller (4), Hartshorn (5), and Colebrook (6). Since Miller was one of the first to point out the importance of inter-electrode

capacities, the effect due to them is sometimes referred to as "*Miller Effect.*" Most of the important results can be obtained from the simplified treatment given below. To avoid complexity in the mathematical working, it will be assumed that the capacities C_1 and C_3 are included in the impedances Z_g and Z_a respectively, with which they are in parallel. In practice it usually happens that these two capacities have no important effect on the circuit. Furthermore, the equivalent series resistances of the inter-electrode condensers will be neglected. Then the circuit of Fig. 4 (*a*) will be equivalent to that

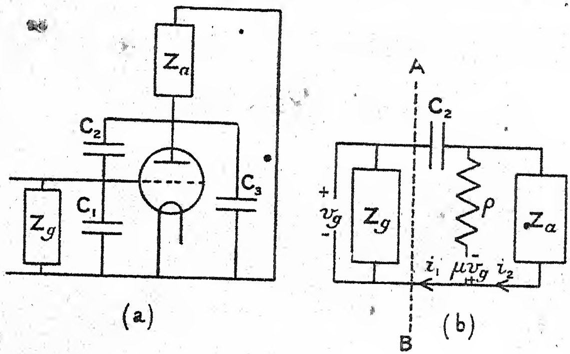

FIG. 4.

shown in Fig. 4 (*b*) where μ and ρ are, respectively, the amplification factor and the slope resistance of the triode, and where it is assumed that a sinusoidal E.M.F. of magnitude v_g is maintained between grid and filament. There are two quantities which it is of special interest to determine. The first of these is the effective impedance of the circuit lying to the right of the dotted line AB (Fig. 4 (*b*)). This may be termed the *Input Impedance* of the triode, and will be denoted by Z_i; it is in parallel with the external impedance Z_g, and with the notation of the figure,

$$Z_i = v_g/i_1.$$

The second quantity, which will be denoted by M, is defined by the equation

$$\mathrm{M} = \mathrm{Z}_a i_2 / v_g \quad . \quad . \quad . \quad (9)$$

The quantity $\mathrm{Z}_a i_2$ is the output potential due to the input potential v_g. Consequently M is the *Voltage Amplification* produced by the valve and its associated circuits. Obviously this is the quantity which will be of importance in the amplifying stages of a receiver.

With the notation of Fig. 4 (*b*) the circuit equations become

$$i_1\left(\rho - \frac{j}{\mathrm{C}_2\omega}\right) - i_2\rho = \left(\mu + 1\right)v_g$$

$$- i_1\rho + i_2(\rho + \mathrm{Z}_a) = - \mu v_g,$$

where ω is 2π times the frequency of v_g. These equations may be solved for i_1 and i_2 in the normal manner. A considerable simplification is introduced by writing

$$r = \frac{\mathrm{Z}_a}{\mathrm{Z}_a + \rho}, \quad g = - \frac{\rho}{j/\mathrm{C}_2\omega} = j\mathrm{C}_2\omega\rho.$$

After simplification of the algebraical expressions, it then appears that

$$\mathrm{Z}_i = - \left(\frac{1 + gr}{1 + \mu r}\right)\frac{j}{\mathrm{C}_2\omega} \quad . \quad . \quad . \quad (10)$$

$$\mathrm{M} = - \left(\frac{1 - g/\mu}{1 + gr}\right)\mu r \quad . \quad . \quad . \quad (11)$$

Before making use of these formulæ, it is necessary to obtain some idea of the magnitudes of the quantities r and g. Replacing Z_a by its equivalent resistance and reactance, we may write

$$r = \frac{\mathrm{R}_a + j\mathrm{X}_a}{\rho + \mathrm{R}_a + j\mathrm{X}_a} \quad . \quad . \quad . \quad (12)$$

Of these terms X_a may theoretically have any value between $-\infty$ and $+\infty$, ρ will always be positive, and it will be assumed that R_a also is always positive. (By the use of special devices, it would be possible to arrange for R_a to have negative values, but this case is at present, of little practical importance.) Consequently, the value of r will always lie between 0 and 1, and will approach the latter limit as either R_a or X_a becomes infinitely large. Turning next to the quantity g and remembering that, for an ordinary triode, C_2 is of the order of $5\mu\mu f$, it is apparent that g will be quite negligible in comparison with unity unless the magnitude of the product $\omega\rho$ be about 10^{11}. Since ρ will never greatly exceed 10^5 ohms, this can only happen when we are dealing with radio frequencies. Thus we may summarize the effects of inter-electrode capacities at *audio frequencies* as follows :

(*a*) The input impedance is equivalent to a capacity, the value of which is C_2 when the anode impedance is zero. The capacity increases with the anode impedance and reaches the limiting value of $(\mu + 1)C_2$, when the anode impedance becomes infinite. This effective capacity is in parallel with the actual grid-filament capacity C_1.

(*b*) The voltage amplification is not affected by the inter-electrode capacities, and is equal to $-\dfrac{\mu Z_a}{Z_a + \rho}$.

At radio frequencies, the quantity g is no longer negligible. When the values of Z_i and M are calculated, it must be remembered that r and g are both " complex " quantities. Let us therefore write

$$r = A + jB, \quad g = jD,$$

where A, B and D are real numbers. Then $D = C_2\omega\rho$ and, for normal triodes, might have a value as large as 1 at a wave-length of 200 metres (i.e. at a frequency of 1500 kilocycles). Also

$$r = A + jB = \frac{R_a + jX_a}{\rho + R_a + jX_a}.$$

From which it follows that

$$A = \frac{R_a(R_a + \rho) + X_a^2}{(R_a + \rho)^2 + X_a^2},$$

$$B = \frac{X_a \rho}{(R_a + \rho)^2 + X_a^2}.$$

Then, since R_a is assumed to be positive, while X_a may have any value, positive or negative, it is clear that A must lie between 0 and 1 and B between $-\frac{1}{2}$ and $+\frac{1}{2}$. When discussing the input impedance at radio frequencies, it is convenient to consider it as a resistance R_i in *parallel* with the capacity C_i. Then, from equation (10) we have

$$\frac{1}{Z_i} = \frac{1}{R_i} + jC_i\omega = \frac{1 + \mu(A + jB)}{1 + jD(A + jB)} \cdot jC_2\omega,$$

from which it follows that

$$R_i = \frac{(1 - DB)^2 + D^2A^2}{\mu D(A^2 + B^2) + DA - \mu B} \cdot \frac{1}{C_2\omega},$$

$$C_i = \frac{(1 - DB) + \mu A}{(1 - DB)^2 + D^2A^2} \cdot C_2.$$

The general variation of these quantities with Z_a is somewhat complicated, but three cases are of particular interest, viz. when the anode impedance is a pure resistance ($X_a = 0 = B$), a pure inductance ($R_a = 0$, X_a positive) and a pure capacity ($R_a = 0$, X_a negative). The variations with A of R_i and C_i for these three cases are shown in Fig. 5 (*a*) and (*b*) which are taken from Colebrook's original article (6). The curves have been calculated for $\omega = 5 \times 10^6$, $D = 0{\cdot}75$; for the sake of convenience the reciprocal of R_i has been plotted rather than R_i itself. When both resistance and reactance are present, the representative curves will lie between those for pure resistance and pure reactance respectively, except at wave-lengths below about 100 metres, when they may lie outside the latter. A general summary of

the mode of variation of the input impedance at radio frequencies is given below.

(*a*) The input capacity depends upon the anode impedance and varies in magnitude from C_2 to a value slightly greater than $(1 + \mu)C_2$. It is greater for inductive than for capacitative impedances, and reaches its maximum value for a *finite* anode inductance.

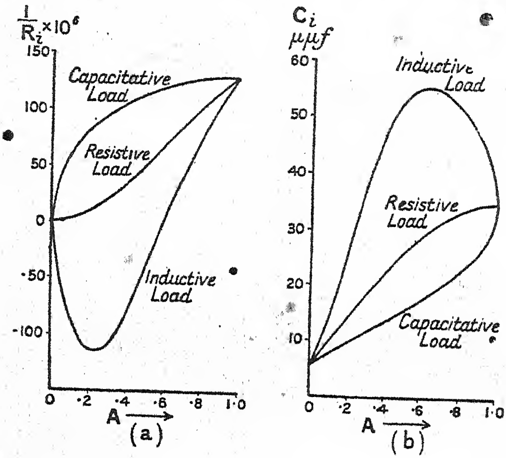

FIG. 5.

Figure 5 is reproduced from "Experimental Wireless," September, 1929, by kind permission of the Editor.

(*b*) The input resistance may be positive or negative for inductive anode impedances, but it is always positive for capacitative impedances. The smallest negative value is attained for a finite anode inductance, and the smallest positive value for an infinite anode impedance

There remains for discussion the expression for voltage amplification at radio frequencies. This will be dealt with in the chapter on High-Frequency Amplification.

CHAPTER III

THE AERIAL-EARTH SYSTEM

WHEN an electromagnetic wave impinges on the aerial system of a wireless receiver, an alternating E.M.F. is set up between two points of this system. Now, when an electromagnetic wave passes through a point in space, there are produced at that point an alternating electric field and an alternating magnetic field, the directions of these fields being at right angles to each other and to the direction of propagation of the wave. The waves used in wireless transmission are usually plane polarized in such a way that the electric vector is vertical and the magnetic vector horizontal. Furthermore, under normal conditions, the electric and magnetic fields are in phase. A discussion of the action of the wave on the aerial system is outside the scope of this book, but one proposition which is generally accepted may be stated as follows. The electric and magnetic fields are merely different manifestations of the wave and each is essential to the wave. They must not be regarded as entirely separate entities, and when it is desired to calculate the E.M.F. set up in an aerial by the wave, this may be done by considering either the electric field *or* the magnetic field, but *not* by considering both and adding the results. Furthermore, for any given aerial system, it is quite immaterial which field we consider, since the calculated E.M.F. will be exactly the same in the two cases. We therefore choose the one which, mathematically, is the more convenient.

Types of Aerial

If we bear the above facts in mind, two distinct types of aerial immediately suggest themselves. Considering only the magnetic field, the most obvious aerial is a plane loop or coil of wire set so that the plane of the wire is normal to the direction of the magnetic field. For a given amplitude of field variation, the E.M.F. set up between the ends of the wire will be proportional to the area of the coil and to the number of turns up to a certain limit. This limit is fixed by the fact that, owing to its inductance and self-capacity, the aerial has a natural frequency of oscillation and if this frequency be less than that of the field variation, standing waves will be set up in the aerial, which can no longer be considered as a simple coil. Aerials of the type described above are known as *Frame Aerials*, and, since they can be made reasonably compact, are frequently used when space is a consideration. In practice the area and number of turns are adjusted so that the natural frequency of the aerial is considerably higher than that of the signal which is to be received. It is obvious that, when a frame aerial is set with its plane parallel to the direction of the magnetic field of a wave, no E.M.F. will be induced in the aerial. This directional property is frequently useful in discriminating between wanted and unwanted signals.

If we consider the electric rather than the magnetic field of an electromagnetic wave, and remember that the direction of the former is vertical, the type of aerial which suggests itself is somewhat as shown in Fig. 6, where A and B are two conductors, separated from each other by a considerable vertical distance. Since one of the requirements of an aerial system is that the two points, between which the E.M.F. is generated, shall be close together, leads are taken from the two conductors A and B to adjacent points P and Q. Then the only stipulation which need be made with regard to the form of A and B is that their electrostatic capacities shall be large compared with those of the connecting wires AP

and BQ. Aerials of a type similar to the one just described were used by Hertz in his original experiments on electromagnetic waves, but it was later found more convenient to dispense with the lower conductor B and connect the point Q to earth. The theory of this is that, assuming the earth to be a perfect conductor, an electric image of A will be produced and will take the place of B. In practice the conductor A usually takes the form of one or more horizontal wires insulated from earth at the ends remote from the vertical connecting wire. The E.M.F. generated in such an aerial will increase both with the height and with the capacity of the top portion, up to a certain limit which, as in the case of a frame aerial, is fixed by the fact that, if the aerial be too large, its natural frequency will be lower than that of the signal and standing waves will be set up. Except in the case of reception of short waves (wave-length less than 100 metres), the limit to the size of an aerial is set rather by practical difficulties of construction than by the natural frequency of the aerial. Some interesting experiments on the design of aerials of this type have been carried out by Smith-Rose and Colebrook (7). Working on a wave-length of about 300 metres, they investigated the manner in which the induced E.M.F. varies with the height and with the form of the top portion of the aerial. They found that, if the top portion consist of two parallel horizontal wires, the E.M.F. is only slightly greater than when a single wire is used, and to obtain this slight advantage, the wires must be about 8 feet apart. The addition of still more parallel wires gives no further improvement. Using a single wire at a height of 25 feet, it is advantageous to increase the length of the top portion up to about 90 feet, but not much beyond this point. Keeping the length of the top portion constant, an increase in height was advantageous up to 30 feet,

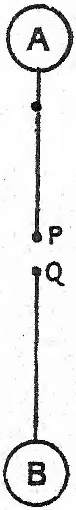

FIG. 6.

at which point the experiments were discontinued. It is reasonable to suppose that similar results would be obtained at other wave-lengths, provided the natural frequency of the aerial were large compared with the signal frequency. This second type of aerial will be referred to as an *Outdoor Aerial*.

Resistance of an Aerial

It can be shown both theoretically and experimentally that an aerial system can be represented by the equivalent circuit shown in Fig. 7, where A and B are the points to which the receiver is connected and E is the

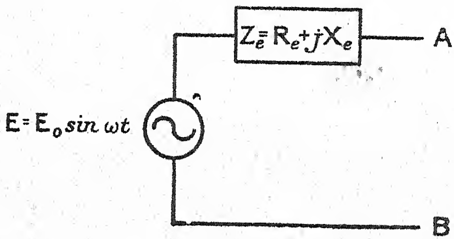

Fig. 7.

E.M.F. induced in the aerial by the signal. Provided the natural frequency of the aerial be greater than the signal frequency, the reactance X will be negative for an outdoor aerial and positive for a frame aerial. It will vary with frequency. For a frame aerial the resistance R will be determined by the length, size and spacing of the wire. In the case of an outdoor aerial, the greater part of the resistance R is located in the earth connection. A water pipe or buried metal plate is generally used for this connection, and, at first sight, either would appear to be an efficient device. However, contrary to what has been assumed above, the earth is not a perfect conductor, and, bearing in mind the electric image theory of the aerial, it is clear that resistance losses will occur in

currents flowing through the earth in the vicinity of the aerial. These losses could be avoided if a perfectly conducting metal plate were placed just above the earth and insulated from it, and if the aerial system were connected to this plate instead of to earth. Such a scheme is hardly practicable, but T. L. Eckersley has shown (8) that the same result may be achieved by using a number of wires parallel to the top portion of the aerial. These wires are placed a few feet above the ground and carefully insulated from it. Such a system of wires is known as an *Earth Screen*. This device was originally developed by Eckersley for use with transmitting aerials, but Smith-Rose and Colebrook have shown that it is equally effective in the case of receiving aerials. Furthermore, they found that little is to be gained in this case by using more than two wires placed about 30 feet apart and 3 feet above the earth. The wires should be rather longer than the horizontal portion of the aerial, and should be symmetrically placed with respect to this portion. With a screen of this type the resistance of an aerial should be only a few ohms, while, with a poor earth connection, it may be as high as fifty ohms.

Selectivity of an Aerial System

The practice of "tuning" an aerial, originally introduced by Lodge, is now almost universally adopted. By "tuning" is meant the addition to an aerial system of such inductances or capacities or both as will cause the system to resonate to the incoming signal. The advantages of this are twofold, since the E.M.F. of the wanted signal is greatly magnified, while the E.M.F. of a signal on any other wave-length is practically unchanged. Consequently, the receiver becomes selective. The *selectivity* of a receiver is a general term used to denote the capability of a receiver to discriminate between two signals on wave-lengths which are not greatly different. Since a receiver will normally be required to operate on

any wave-length within a prescribed range, it follows that at least one of the aerial circuit elements must be continuously variable over a definite range of values, to enable the aerial to be tuned to the required frequency. While either a capacity or an inductance may be made continuously variable, the former plan is much the more convenient in practice, and is usually adopted.

Before discussing the methods which may be used to tune an aerial, we must consider further the question of selectivity. In this connection it must be remembered that the incoming signal will, in general, comprise a carrier wave and two side-bands. Assuming that the

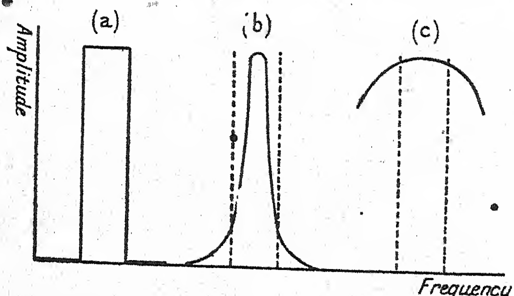

Fig. 8.

highest note which it is desired to reproduce has a frequency of 10 kilocycles, the frequencies of the side-bands may lie anywhere within a range of 10 kilocycles on either side of the frequency of the carrier wave. To avoid distortion, it is necessary that the relative amplitudes of carrier wave and side-bands should be unchanged by resonance in the aerial circuit. Consequently, the ideal shape for the resonance curve of the aerial circuit would be as shown at (*a*) Fig. 8. If the curve be of the form represented by (*b*), which might be obtained with a low resistance tuned circuit, the output amplitude of a side-band for a given input amplitude will decrease as the frequency difference between carrier wave

and side-band increases. Now, this frequency difference is equal to the frequency of the note which is being transmitted. It therefore follows that the use of a circuit with a resonance curve such as (*b*) will cause a reduction of the amplitudes of high notes in the speech or music which is being transmitted. This effect, which is a particular type of frequency distortion, is known as *Side-Band Cut-off*. It can be avoided by increasing the resistance of the tuned circuit so that the resonance curve takes the form shown in (*c*) Fig. 8. If this be done, however, the selectivity of the receiver will be so poor that it will be difficult to separate transmissions from two stations on neighbouring wave-lengths. Thus it appears that a suitable resonance curve cannot be obtained by using a single-tuned circuit. We shall return to this point later, when the various stages of a receiver have been considered.

General Theory of Aerial Tuners

The methods which may be used for tuning an aerial are so numerous that it will be impossible to consider each one in detail. However, a great deal of information concerning the principles involved may be gathered from a study of the circuit shown in Fig. 9, where E is the E.M.F. induced in the aerial and Z_e is the combined impedance of the aerial and any circuit element which may, for any purpose, be added to the aerial. The points P and Q are supposed connected between grid and filament of the first valve of the High-Frequency Stage and therefore, in accordance with the principles laid down in Chapter II, the impedance between these two points will depend upon the load in the anode circuit of this valve, and may be represented by a resistance R_o shunted by a capacity. For reasons which will appear later, we suppose an additional variable reactance to be placed in parallel with R_o, so that the total impedance Z_o, between P and Q may be represented as shown in the figure, by a resistance R_o shunted by a

reactance X_o. The reactances which are added to the aerial circuit will, in practice, of necessity be accompanied by resistances. For the sake of simplicity we suppose these resistances included in the terms R_e and R_o. Let the E.M.F. E give rise to a voltage V between the points P and Q; then V is the input E.M.F. to the High-Frequency Stage, and it is obviously desirable that this E.M.F. should be as large as possible. We proceed to investigate the effect of variation of the circuit elements

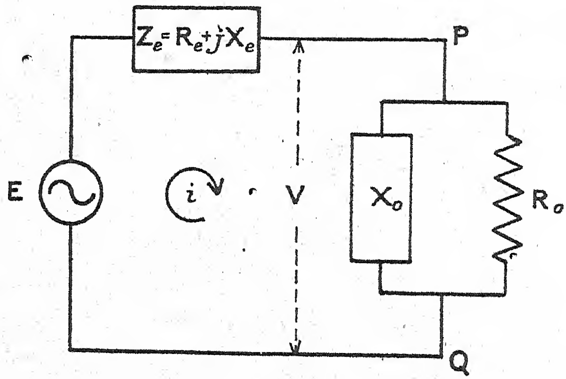

FIG. 9.

on (*a*) the ratio V/E, and (*b*) the selectivity of the circuit. With the notation of the figure, we may write

$$\frac{1}{Z_o} = \frac{1}{R_o} + \frac{1}{jX_o},$$

$$Z_o = jR_oX_o/(R_o + jX_o) = R_1 + jX_1 \text{ (say)}.$$

Then $|Z_o|^2 = R_o^2X_o^2/(R_o^2 + X_o^2)$

and $R_1 = R_oX_o^2/(R_o^2 + X_o^2) = |Z_o|^2/R_o$. (13)

Furthermore $i = E/(Z_e + Z_o)$.

Now for a given value of Z_o, V will be a maximum when i is a maximum, and this will occur when $X_e = -X_1$;

i.e. when the circuit is tuned to resonance. Assuming this condition to be fulfilled,

$$V = Z_o i = EZ_o/(R_e + R_1),$$

so that $\quad V/E = Z_o/(R_e + R_1).$

Considering only the absolute magnitude of V/E without regard to phase, we may write

$$|V/E| = |Z_o|/(R_e + R_1),$$

or substituting from (13) for R_1

$$|V/E| = R_o|Z_o|/(R_oR_e + |Z_o|^2).$$

By differentiation we can show that this expression has a maximum value when

$$R_oR_e = |Z_o|^2 \quad . \quad . \quad . \quad (14)$$

so that, finally, when this condition is fulfilled,

$$|V/E| = \tfrac{1}{2}\sqrt{R_o/R_e} \quad . \quad . \quad . \quad (15)$$

We may further note that, in this case

$$R_1 = |Z_o|^2/R_o = R_e,$$

so that the equivalent *series* input resistance of the valve is equal to the resistance of the aerial.

From the above analysis it becomes clear that, in order to obtain the maximum value of V, *two* adjustments of aerial reactances are necessary. The first of these establishes equality between the aerial resistance and the equivalent series input resistance to the High-Frequency Stage, while the second brings the circuit to resonance. Now, of these two, the second is vastly more important than the first, since failure to establish resonance may decrease the value of V by a factor as high as 50, while incorrect matching of the two resistances will seldom reduce the value of V by more than about 30 per cent. While, therefore, one continuously variable reactance is essential to an aerial circuit to

allow tuning over a range of wave-lengths, the second reactance is, for the sake of simplicity, usually fixed at some definite value and the matching of resistances is then not accurately carried out.

Factors Affecting Selectivity

Consider next the question of selectivity. The current flowing round the circuit (Fig. 9) is given by

$$|i| = E/[(R_e + R_1)^2 + (X_e + X_1)^2]^{\frac{1}{2}}.$$

Since the reactances X_e and X_1 must be of opposite sign for resonance to be obtained, we will suppose one of them to be an inductance L and the other a capacity C. Then

$$|i| = E/\sqrt{(R_e + R_1)^2 + (L\omega - 1/C\omega)^2}$$

and the circuit will resonate when

$$\omega = \omega_o = 1/\sqrt{LC},$$

giving

$$|i|_{\omega_o} = E/(R_e + R_1).$$

If now the frequency be changed to some neighbouring value so that ω becomes $\omega_o + \delta\omega$, where $\delta\omega$ is small compared with ω_o, the current will be given by

$$\begin{aligned} |i|_{\omega_o+\delta\omega} &\\ &= E/\sqrt{(R_e + R_1)^2 + \left[L(\omega_o + \delta\omega) - \frac{1}{C\omega_o}\left(1 - \frac{\delta\omega}{\omega_o}\right)\right]^2} \\ &= E/\sqrt{(R_e + R_1)^2 + 4L^2(\delta\omega)^2}. \qquad (16) \end{aligned}$$

Then the ratio $|i|_{\omega o}/|i|_{\omega o+\delta\omega}$ will give an indication of the selectivity of the circuit, and from equation (16) we see that the conditions for this ratio to be large are :

(*a*) R_e and R_1 to be as small as possible.
(*b*) With a given value of $R_e + R_1$, the ratio L/C to be as large as possible.

The second condition is usually of minor importance, since the ratio L/C is fixed by considerations of the band of wave-lengths to be received. Of the quantities R_e and R_1, the former will not greatly exceed the resistance of the aerial, provided properly designed reactances be used, and so cannot be varied to any great extent. Thus the selectivity of the circuit is determined largely by the value of R_1 which can be adjusted by suitable choice of X_o. Obviously, the value of X_o which gives the required selectivity will not, in general, be the same as that which secures matching of R_e and R_1, but usually selectivity is the more important factor.

The above theory has been worked out with reference to the circuit of Fig. 9, and it will be appreciated that all practical aerial circuits do not agree in detail with this particular one. Nevertheless, the underlying principles are always the same and there should be no difficulty in modifying the theory to suit any given case.

Types of Aerial Tuner

By way of illustration of the above theory, we give details of some of the more commonly used aerial tuning

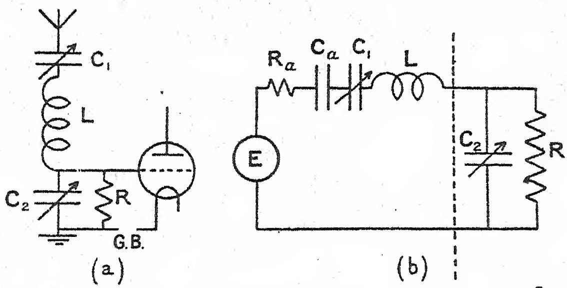

Fig. 10.

circuits. Consider first the simple Series-tuned circuit represented in Fig. 10 (a). The equivalent circuit is drawn in Fig. 10 (b), and corresponds exactly to the case

considered above, the elements to the left and right of the dotted line forming the impedances Z_e and Z_o respectively. The more usual Parallel-tuned circuit is illustrated in Fig. 11 (*a*) and (*b*), where the dotted line is

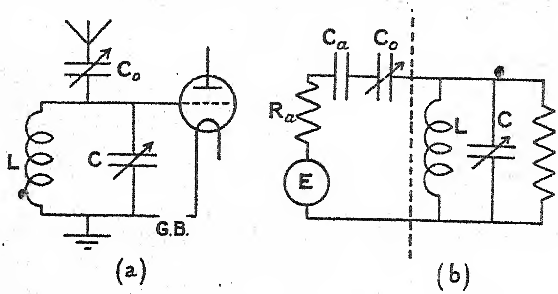

FIG. 11.

drawn as before to mark the boundary between the impedances Z_e and Z_o. The capacity C_o is frequently fixed, and is usually given a value less than the optimum one in order to improve the selectivity of the circuit.

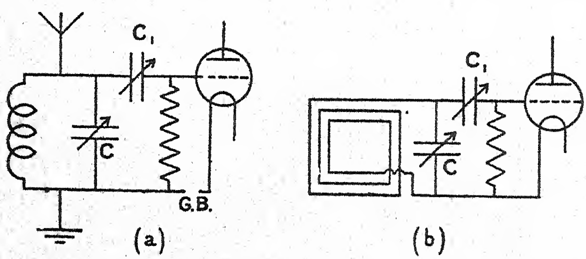

FIG. 12.

Sometimes it is omitted altogether when selectivity is not an important factor. It may be noticed that, if the capacity C_o be very much smaller than C_a, the effective capacity of C_a and C_o in series will be practically equal

to C_o. The setting of condenser C to tune to any particular wave-length will then be independent of the aerial to which the receiver is attached. This fact is of great commercial importance, since if the receiver be calibrated during course of manufacture, the calibration will be unaltered when the receiver is attached to a different aerial.

Another type of aerial system is shown in Fig. 12,* where the applications to an outdoor aerial (a) and a frame aerial (b) are illustrated. Although the details are not quite the same as in the circuit of Fig. 9, the principles are identical in the two cases. The circuit is tuned by adjustment of C and the selectivity is controlled by means of C_1.

Equivalent Circuit of a Transformer

We turn now to a rather different type of aerial tuning system, viz. that in which a transformer is used. Since

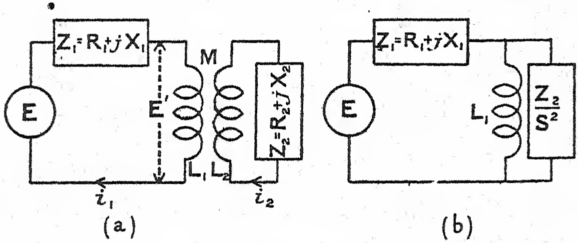

Fig. 13.

transformers are of considerable importance in radio receivers, we proceed to establish some of their more important properties. Consider the circuit of Fig. 13 (a) where an E.M.F. E in series with an impedance $Z_1 = R_1 + jX_1$ is applied to the primary of a transformer

*In Fig. 12 (b) and in many of the circuit diagrams which follow, the Grid Bias battery (G.B.) has for simplicity been omitted.

the secondary of which is connected to an impedance $Z_2 = R_2 + jX_2$. Let the primary, secondary, and mutual inductances be L_1, L_2, and M respectively, and let the resistances of the windings be included in R_1 and R_2. It will be assumed that there is perfect coupling between the two windings, so that

$$M^2 = L_1 L_2.$$

If we further assume that L_1 and L_2 are proportional to the squares of the numbers of turns on primary and secondary respectively, we may write $S = \sqrt{L_2/L_1}$ where S is the turns ratio of the transformer.

Let E′ be the E.M.F. across the primary winding after allowing for the voltage drop in the impedance Z_1. Then, with the notation of the figure,

$$jL_1\omega i_1 + jM\omega i_2 = E'$$
$$jL_2\omega i_2 + jX_2 i_2 + R_2 i_2 + jM\omega i_1 = 0.$$

From which it can be shown that

$$i_1 = \frac{E'(R_2 + j_2X + jL_2\omega)}{jR_2L_1\omega - L_1X_2\omega} = E'\left\{\frac{1}{jL_1\omega} + \frac{L_2/L_1}{R_2 + jX_2}\right\} \quad (17)$$

$$i_2 = \frac{E' - jL_1\omega i_1}{jM\omega} = \frac{E'\sqrt{L_2/L_1}}{R_2 + jX_2} \quad . \quad . \quad . \quad . \quad (18)$$

From equation (17) we see that, so far as the primary current is concerned, the circuit of Fig. 13 (*a*) is equivalent to that of Fig. 13 (*b*), where the secondary of the transformer has been removed and the secondary load impedance divided by S^2, has been placed in parallel with the primary winding. Furthermore, it is clear from equation (18) that the E.M.F. across the secondary winding is S times the E.M.F. E′ across the primary winding.

The circuit of Fig. 14 (*a*) provides an example of the application of a transformer to an aerial tuning system. R represents the input resistance of the first valve of the

receiver. In this case a "tapped" coil is used to form an auto-transformer, but the theory developed above still applies except that the coupling between primary and secondary will not be perfect, and equation (17) will not hold. However, it can be shown that this alteration will make very little difference to the results previously obtained, provided the coupling be not too loose. The equivalent circuit is shown in Fig. 14 (*b*). Resonance is obtained by variation of the capacity C and matching of the resistances by correct choice of the transformer ratio S. The input E.M.F. to the first valve will be S times the E.M.F. E′ across the inductance L. If the

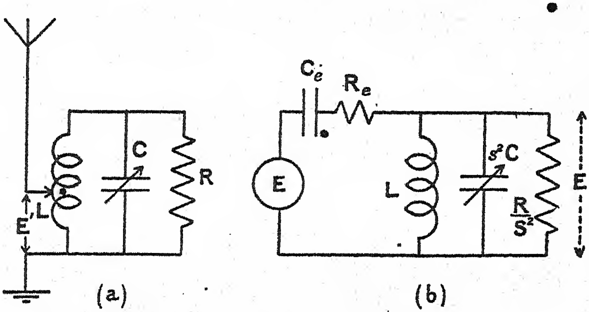

FIG. 14.

transformer ratio be high, the frequency calibration of the condenser C for resonance will be almost independent of the constants of the aerial.

STABILITY OF AN AERIAL SYSTEM

In conclusion, we must consider what will happen if the input resistance of the H.F. amplifier should become negative. In this case, it is obvious that the quantity of energy supplied from the aerial to be dissipated by the H.F. amplifier is negative. This is equivalent to saying that energy is flowing from the H.F. amplifier into the aerial system. If the rate at which this flow takes place

be greater than the rate at which the energy can be dissipated by resistances in the aerial circuit, the E.M.F. in the aerial circuit will increase indefinitely. It follows that the slightest electrical disturbance of any kind in the aerial circuit, whether due to a signal or not, would cause this circuit to burst into oscillation. Furthermore, the rate at which energy is fed into the aerial circuit will be greater for a small value of negative resistance than for a large value. Thus it follows that, for any given aerial circuit, there will be a critical value of resistance; if the input resistance of the H.F. amplifier be negative and less than this critical value, instability will result. This conclusion is entirely confirmed by a more rigorous mathematical analysis, and by the experimental investigations of Miller (4).

CHAPTER IV

HIGH-FREQUENCY AMPLIFICATION

THE function of the High-Frequency stage of a wireless receiver is to magnify the H.F. potential which is applied to it by the aerial tuning system. The general circuit of a H.F. amplifier is indicated in Fig. 15, where the input E.M.F. e_g is applied between grid and filament of the valve V_1. Included between the anode of this valve and the positive end of the H.T. supply is an impedance Z_a, which must provide a path for the steady current necessary for the operation of V_1. Let e_g' be the E.M.F. produced between grid and filament of the valve V_2; then, since we may regard e_g' as the output E.M.F. from V_1, the ratio e_g'/e_g measures the amplification due to the stage which includes V_1 and the associated circuit elements. Theoretically, any number of stages may be placed in cascade for the purpose of obtaining greater amplification, but the practical difficulties become very great if more than two or three stages be used. The underlying principles remain the same, however many stages are used, so it will be unnecessary to consider more than one.

Whether the valve V_2 form part of a second H.F. stage or a part of the Detector stage, it will require a steady grid potential considerably different from that of the point A, Fig. 15. A condenser C is therefore inserted between the grid and this point to prevent the flow of direct current while, at the same time offering a free path to the flow of high-frequency alternating current.

A common value for the capacity of this condenser is about 0·0003 μf, and at a frequency of a million cycles per second it would then have a reactance of about 500 ohms. A resistance R, known as a *Grid Leak*, in series with a grid bias battery (G.B.) connects the grid and filament of V_2, and by this means the steady potential of the grid is maintained at an appropriate value. Similarly the steady potential of the grid of V_1 is maintained

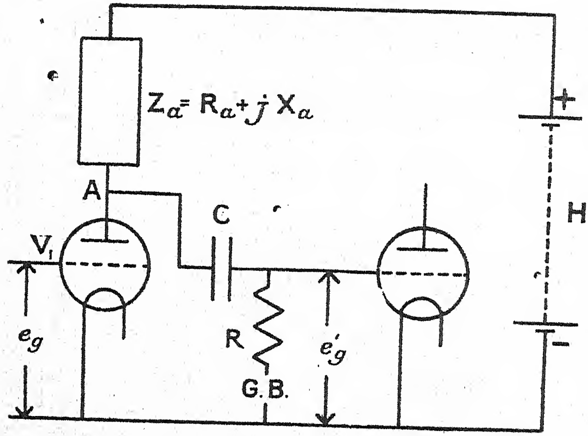

FIG. 15.

at a suitable value by a battery included in the aerial tuning system.

The choice of this potential in the case of a H.F. amplifier is limited by two considerations. In the first place it should be such that the grid potential variations due to the input E.M.F. e_g never cause the operating point to stray from the straight portions of the valve characteristics (see p. 11). Secondly, the grid must never acquire a positive potential, since this would cause grid current to flow, thus adding a variable load to the aerial tuning system. The precise effects of this load

are not easy to calculate, but it would certainly cause loss of signal strength and probably distortion also.

Before considering the nature of the impedance Z_a, it will be convenient to re-draw the circuit of Fig. 15, replacing V_1 by its equivalent circuit and omitting all sources of steady potential. This has been done in Fig. 16, where R_i and C_i are respectively the input resistance and capacity of V_2, ρ is the slope resistance, and μ the amplification factor of V_1. Interelectrode capacities of V_1 are for the moment neglected; their importance lies in the effect which they have upon the

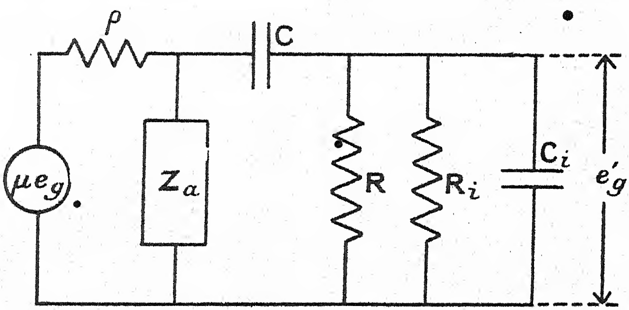

FIG. 16.

input resistance of the H.F. stage, and this will be considered in detail later on.

In practice the reactance of C (Fig. 16) will be negligibly small compared with the impedances of the other circuit elements, so we may omit it. Then writing Z for the impedance of Z_a, R, R_i and C_i in parallel,

$$\text{Voltage Amplification} = e_g'/e_g = \mu Z/(Z + \rho).$$

Thus, in order that the amplification shall approach the limiting value μ, it is necessary that Z be large compared with ρ. Now ρ will not usually be less than 10,000 ohms, so that Z should be at least 50,000 ohms. Of the elements which are included in Z, R can be made so

large that it will not appreciably affect the value of Z, in practice it is usually of the order of one megohm. The resistance of R_i will depend upon the load in the anode circuit of V_2, but will usually be an appreciable fraction of a megohm and will not greatly affect the value of Z.

Resistance Coupled Amplifiers

Consider first the case of *Resistance Coupling*, when Z_a is a pure resistance of value R_a. Then the absolute magnitude of Z is given by

$$|Z| = R_a/\sqrt{1 + R_a^2C_i^2\omega^2},$$

where $\omega = 2\pi f$. Consequently, the maximum possible value of $|Z|$ is $1/C_i\omega$, which will occur when R_a is infinitely great. Now, as explained in Chapter II, the value of C_i will depend chiefly on the anode-grid capacity of V_2 and the impedance in the anode circuit of V_2, but it will not usually be much less than 50 $\mu\mu f$, so that, at a frequency of one million cycles per second, the limiting value of $|Z|$ will be about 3000 ohms, and this is too small to give efficient amplification. On account of this shunting effect of the capacity C_i, resistance coupling is not generally used except on wave-lengths greater than 1000 metres, where the reactance is correspondingly higher.

However, special valves have been placed on the market in which the above-mentioned difficulty has to some extent been overcome. A diagrammatic sketch of a valve of this type is given in Fig. 17, from which it will be seen that two triodes, together with the necessary resistance coupling elements, are all enclosed in the same glass envelope. By this mode of construction inter-electrode capacities are reduced to a minimum and an amplification of about four times may be obtained from the triode V_1 at a frequency of a million cycles per second. Resistance or any other type of coupling may

be employed between V_2 and the valve which follows it. For further details of multiple valves of this type the reader is referred to the bibliography at the end of the book (9).

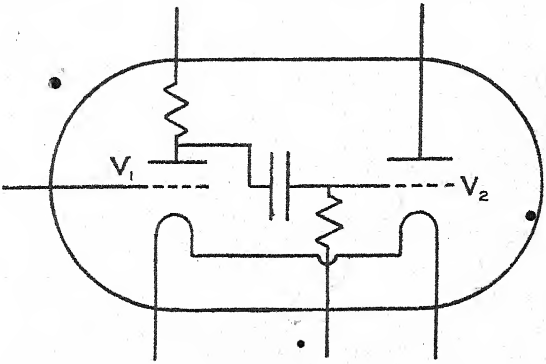

FIG. 17.

CHOKE COUPLING

Consider next the case where the impedance Z_a (Fig. 16) is a High-Frequency Choke. The stage is then said to be *Choke Coupled* and the circuit is as shown in Fig. 18, where the capacity C_1 includes both the self-capacity of the choke and the input capacity of the next valve, while the resistance R_1 includes the input resistance of this valve, the equivalent shunt resistance of the choke itself and the resistance of the grid leak R (Fig. 16). Let Z be the impedance of L, C_1 and R_1 in parallel. Then

$$\frac{1}{Z} = \frac{1}{R_1} + \frac{1}{jL\omega} - \frac{C_1\omega}{j} = \frac{1}{R_1} + j(C_1\omega - 1/L\omega) \quad (19)$$

Let g be the voltage amplification due to the stage; then

$$g = \mu Z/(Z + \rho) = \mu/(1 + \rho/Z)$$
$$= \mu/\left\{1 + \frac{\rho}{R_1} + j\rho(C_1\omega - 1/L\omega)\right\} \quad . \quad (20)$$

Thus g will be a maximum for that value of ω (say ω_o), for which $C_1\omega = 1/L\omega$. Then we may write

$$g_{\omega_o} = \mu R_1/(R_1 + \rho) \quad . \quad . \quad . \quad (21)$$

Since the amplifier will normally be required to operate efficiently over a band of wave-lengths, it is desirable to design the choke in such a way that the value of g shall be as nearly as possible equal to that given by (21) when ω is not equal to ω_o. In order to do this, we see from equation (20) that, for a given value of ω, the quantity $C_1\omega - 1/L\omega$ must be as small as possible.

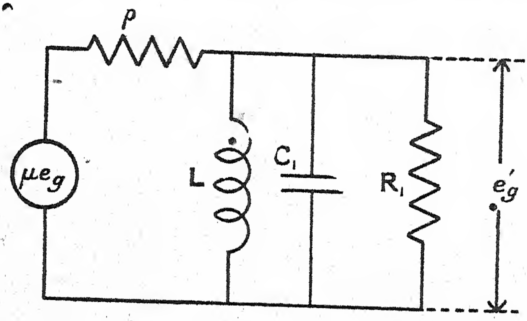

FIG. 18.

Bearing in mind that, for a given value of ω the product LC_1 is fixed, we see that this condition will be satisfied when the ratio C_1/L is as small as possible. This assumes that R_1 is independent of this ratio, whereas, for a choke of given dimensions, a decrease of the C_1/L ratio will usually involve a decrease of the equivalent shunt resistance of the choke.

At first sight it might appear that the choke should be designed in such a manner that its resonant frequency would lie about midway within the band of frequencies to be amplified. This, however, is seldom done, for reasons which we must now consider. It is clear that the choke will behave as a capacity or an inductance,

according as its resonant frequency is less or greater than the frequency of the signal which is being amplified. In the first case the input resistance of the stage will always be positive, but in the second case it may be either positive or negative so that, as explained on page 33, there is a possibility that spontaneous oscillation may result. To avoid this possibility, H.F. chokes are generally designed so that, when connected in the anode circuit of a valve, the resonant frequency of the resulting combination shall be slightly less than the lowest frequency to be amplified.

Both in resistance-coupled and choke-coupled H.F. amplifiers the circuit elements are all fixed and are designed in such a way that the amplification does not vary greatly over a range of frequencies. Such amplifiers are said to be *Aperiodic*.

Transformer Coupling

We turn now to *Tuned* amplifiers in which at least one of the circuit elements is variable so that its magni-

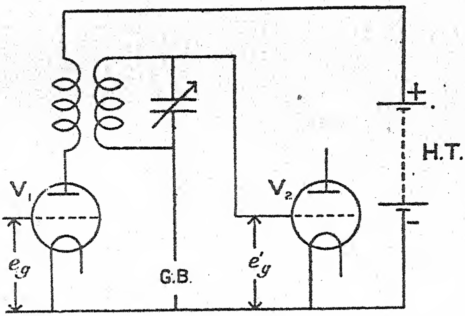

Fig. 19.

tude can be adjusted to suit the particular frequency of the signal which is being amplified. The most general type of tuned amplifier is that in which *Transformer*

Coupling is employed, the circuit being as shown in Fig. 19. For purposes of calculation we may replace this by the equivalent circuit shown in Fig. 20 where μ and ρ are respectively the amplification factor and slope resistance of V_1, L_1 the inductance and r_1 the resistance of the transformer primary, L_2 and r_2 the corresponding quantities for the secondary, M the mutual inductance between primary and secondary, and C_2 a variable capacity by means of which the circuit is tuned to resonance. The input capacity of V_2 is supposed included in C_2 with which it is in parallel. Similarly, the input resistance

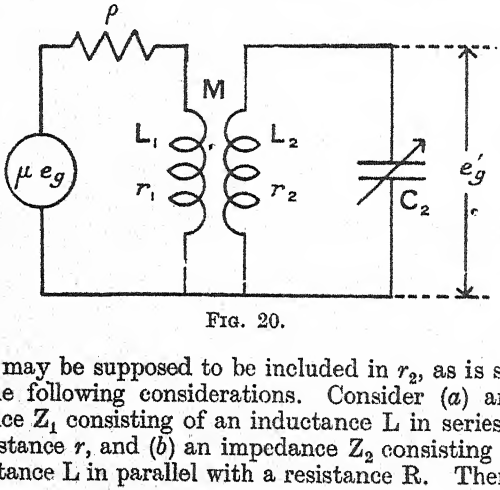

FIG. 20.

of V_2 may be supposed to be included in r_2, as is shown by the following considerations. Consider (*a*) an impedance Z_1 consisting of an inductance L in series with a resistance r, and (*b*) an impedance Z_2 consisting of an inductance L in parallel with a resistance R. Then

$$Z_1 = jL\omega + r$$
$$Z_2 = jL\omega R/(R + jL\omega)$$
$$= (L^2\omega^2 R + jL\omega R^2)/(R^2 + L^2\omega^2).$$

If R be large, so that $L^2\omega^2$ can be neglected in comparison with R^2, we have

$$Z_2 = \frac{L^2\omega^2}{R} + jL\omega.$$

Then the impedances Z_1 and Z_2 will be identical, provided $r = L^2\omega^2/R$. R is frequently termed the "dynamic" resistance of the tuned circuit. Hence in Fig. 20, the resistance of the transformer secondary and the input resistance of V_2 may both be represented by a single series resistance r_2 or by a single parallel resistance R_2, where $R_2 = L^2\omega^2/r_2$. Sometimes one and sometimes the other form is more convenient.

The circuit of Fig. 20 has been discussed by several writers, including Browning and Drake (10) and McLachlan (11). The treatment here given is due to McLachlan. Let κ be the coefficient of coupling so that $M^2 = \kappa^2 L_1 L_2$, and let $L_2/L_1 = S^2$, then in general S will be approximately equal to the turns ratio of the transformer. The resistance r_1 will be quite negligible in comparison with ρ and will be omitted. Then solving the circuit equations in the normal manner and writing $\kappa/S = a$, we find

$$e_g' = \frac{L_2 \mu e_g}{aL_2 + \rho r_2 C_2/a} \quad . \quad . \quad . \quad (22)$$

Now the same transformer will usually be required to operate over a range of frequencies and, since resonance must always be established, C_2 will be inversely proportional to the square of the frequency. Furthermore, r_2 will increase with frequency (owing to "skin" effect, etc.), and, over the working range, can be shown to be approximately proportional to frequency. Therefore $r_2\sqrt{C_2}$ will remain approximately constant, while r_2C_2 decreases with frequency. Now from equation (22) it is easy to show that the maximum value of e_g' will result when

$$aL_2 = \rho r_2 C_2/a$$

i.e.

$$a = \sqrt{\rho r_2 C_2/L_2} \quad . \quad . \quad . \quad (23)$$

and, unless the coefficient of coupling of the transformer can be varied, this condition can only be satisfied for one particular frequency. In practice it is usual to

employ fixed coupling and design the transformer for some frequency intermediate between the maximum and minimum values which may be encountered.

When the transformer is tuned to resonance, the impedance of its secondary will be equal to its "dynamic" resistance R_2, where $R_2 = \omega^2 L_2{}^2/r_2 = L_2/r_2C_2$.

Substitution of this in (23) gives us for the optimum ratio,

$$a = \kappa/s = \sqrt{\rho/R_2}$$

and with this value of a

$$e_g' = \tfrac{1}{2}\mu e_g\sqrt{R_2/\rho} \quad . \quad . \quad . \quad (24)$$

Thus for efficient amplification, R_2 must be large, and since the magnitude of R_2 is governed to a large extent by the series resistance of the secondary of the transformer, this resistance should be kept as small as possible. It is not important for the resistance of the transformer primary to be low, since in any case, it will be almost negligible in comparison with the slope resistance of the valve with which it is in series. One important consequence of equation (24) is that the quantity which determines the efficiency of a valve in a H.F. amplifier is the ratio $\mu/\sqrt{\rho}$, so that it is advantageous to use valves with a high value of μ.

The theory given above enables us to calculate the optimum transformer ratio for any given set of conditions, but our final choice of ratio may be influenced by other factors such as the effect which it will have upon the selectivity of the receiver. In estimating this we must remember that the value of C_2 (Fig. 20) is to a large extent limited by conditions of manufacture and the band of wave-lengths to be covered, so that we may regard it as fixed for any one frequency. Then assuming that there is perfect coupling between primary and secondary and transferring the capacity C_2 and resistance r_2 to the primary circuit (as explained on p. 31), we see that the effective capacity there is S^2C_2. Hence an increase in

the transformer ratio will increase the effective capacity of the tuned circuit and a gain in selectivity will result.

Tuned Anode Coupling

A form of coupling which is frequently used is depicted in Fig. 21, and is known as the *Tuned Anode*. We may regard it as a tuned auto-transformer of unit ratio, to which are added a condenser C′ and grid leak R′ in order that the steady potential of the following valve may be maintained at a suitable value. The equivalent circuit

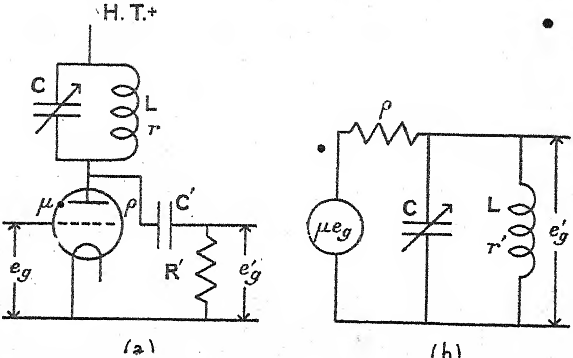

Fig. 21.

is shown in Fig. 21 (*b*), where *r′* includes the added effective resistance due to the grid leak in parallel with the inductance, as well as the true resistance *r* of the inductance; the capacity C′ is omitted, since it will be sufficiently large for its reactance to be negligible in comparison with other impedances in the circuit. The chief advantage of the Tuned Anode is that it enables us to dispense with a double winding on the transformer, and so simplify the construction of the H.F. stage.

Another form of coupling is illustrated in Fig. 22 (*a*)

and, omitting the capacity C′ as before, the equivalent circuit is shown in Fig. 22 (*b*). If the H.F. choke L′, r' be well designed, the load which it adds to the tuned circuit will be very small, so that this form of coupling is essentially the same as the Tuned Anode.

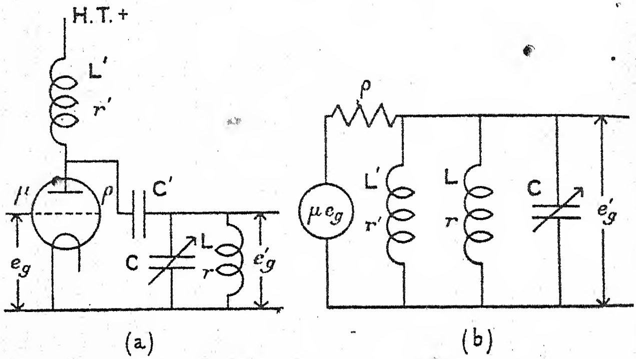

Fig. 22.

Stability of an Amplifier

We must now consider the effect which the anode-grid capacity of the valve of a tuned H.F. amplifier will have upon the input resistance of this stage. We will first determine how the input resistance changes with variation of the capacity C of the tuning condenser of the amplifier, and we will suppose that when C is equal to C_o the amplifier is tuned to the frequency of the incoming signal. Then, bearing in mind that the anode circuit of the amplifier will behave as an inductance, a pure resistance or a capacity according as C is less than, equal to, or greater than C_o, and applying the formulæ derived in Chapter II (cf. Fig. 5, p. 18), we see that the general form of the variation of the input resistance with C will be as shown in Fig. 23 which, however, is not drawn to scale.

It thus appears that, over a certain range of values of C, the input resistance of the stage will be negative and, since the stage will usually be preceded by a tuned aerial circuit, there is a possibility that spontaneous oscillation may occur. Whether it will, in fact, occur will depend upon the constants of the aerial circuit and the smallest negative value attained by the input resistance of the H.F. stage, and this latter, in turn, will depend upon the constants of the anode circuit and the anode-grid capacity of the valve of the H.F. stage. At first sight it might be argued that this effect is of little

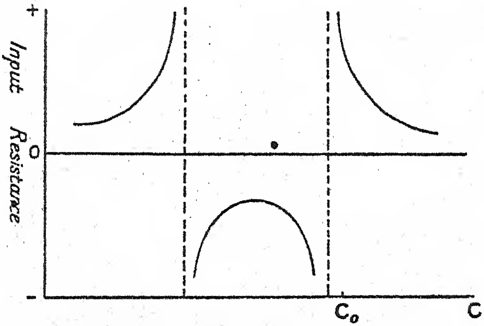

FIG. 23.

importance since, when the H.F. stage is tuned to the signal frequency its input resistance will be positive and therefore spontaneous oscillation cannot occur. This argument, however, is unsound because, although it is true that oscillation cannot occur *at the signal frequency*, it can and often will occur at some slightly lower frequency for which the input resistance is negative.

The mathematical theory of this effect has been worked out by Beatty (12) for Tuned Anode coupling, and by Butterworth (13) for Transformer Coupling. For our present purpose it will be sufficient to state that, in both cases, spontaneous oscillation will always occur

when efficient coils and tuning condensers are used in conjunction with an average triode, so that efficient H.F. amplification cannot be obtained unless some steps be taken to overcome or diminish the effect which we have been considering.

It was shown in Chapter II that the input resistance of a valve, and therefore of a H.F. amplifier stage, depends to a very great extent upon the anode-grid capacity, C_{ag}, of the valve. In particular, if C_{ag} could be made sufficiently small, the absolute value of the input resistance, whether positive or negative, would be so high that there would never be any danger of spontaneous oscillation. Now it must be remembered that the effective anode-grid capacity includes not only the actual internal capacity between the valve electrodes, but also the external stray capacity between all wires connected to these electrodes. Obviously, therefore, the first step lies in eliminating these external stray capacities, and this is always done in modern receivers by surrounding all components by metal screens which act as electrostatic shields. These screens have the further advantage that, if they be properly constructed, the eddy currents induced in them will effectively neutralize any inductive coupling between adjacent components. For further information on this subject the reader is referred to an article by Smith-Rose (14).

When all external stray capacities have been eliminated by screening, the remaining internal capacity of an ordinary triode is still too large for efficient stable H.F. amplification to be obtained. An early attempt to overcome this difficulty was made by Round, who constructed valves in which the leads to the various electrodes were brought through the glass envelope at points well separated from each other. These valves never became very popular, partly because the reduction in capacity was not sufficient and partly because the valves were difficult to construct commercially. The first real solution to the problem was put forward by Hazeltine under the name of the "Neutrodyne" circuit. Several

modifications of the original circuit have since appeared, but the underlying principle is the same in all of them, and consists in arranging the H.F. stage in the form of a "bridge" circuit and adding a small "neutrodyne" capacity in such a way as to neutralize the valve inter-electrode capacity. The arrangement for a "neutralized" tuned anode H.F. stage, together with the equivalent circuit, is illustrated in Fig. 24, where C_{ag} is the anode-grid capacity of the valve and C_2 the capacity of the neutrodyne condenser. We shall not consider these

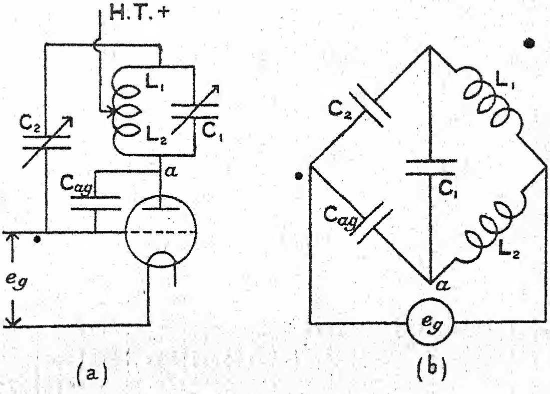

FIG. 24.

interesting circuits in greater detail because, at the present time, they have been rendered almost superfluous by the introduction of the "Screen-Grid" valve.

SCREEN-GRID VALVES

Since the early days of the introduction of the thermionic valve into wireless receivers, experiments have been carried out with valves containing more than three electrodes. One of the earliest workers in this field was

Schottky, who used valves in which two grids were interposed between the filament and anode. He realized quite clearly that, if the outer of these grids were maintained at a definite steady potential with respect to the filament, it would act as an electrostatic screen between the inner grid and the anode without seriously impeding the flow of electrons from filament to anode. Schottky does not seem to have made much use of this idea, and it was not until 1925 that Hull (15) showed that the arrangement might be successfully used to overcome the harmful effects of anode-grid capacity in H.F. amplifiers. It will be convenient to speak of the inner grid as the *control* grid, and the outer grid as the *screen* grid. Hull constructed valves in which the electrostatic screening was as complete as possible, and thus he was able to design H.F. amplifiers which were very much more efficient than any previously used. However, his valves were somewhat difficult to construct and did not become very popular. A year or two later Round designed Screened-Grid valves which could be manufactured commercially, and since that time similar valves have been used almost universally in H.F. amplifiers. In the modern Screened-Grid valve the anode lead is brought to a separate terminal on the top of the bulb, while the remaining leads are taken to a cap on the base. The potential of the screen grid is maintained at some value intermediate between those of the filament and plate respectively. The screen grid is usually constructed in such a manner that it has a flange extending to within a millimetre or two of the walls of the bulb so that the shielding effect may be continued almost without break by an external metallic screen. When this external screen is used the residual anode-grid capacity may be of the order of 0·005 $\mu\mu f$, or about one-thousandth part of that which obtains in an ordinary triode.

We have seen above that the screen-grid valve was originally introduced to overcome the harmful effects due to anode-grid capacity in H.F. amplifiers. The introduction of the screen grid has, however, an addi-

tional very important effect upon the characteristics of the valve in that it materially weakens the control which the anode is able to exert on electrons near the filament. As a result, both the slope resistance and amplification factor of the valve are greatly increased, and since, as we have seen, the efficiency of the valve in a H.F. amplifier depends upon the ratio $\mu/\sqrt{\rho}$, the introduction of the screen-grid is beneficial, quite apart from its use in preventing spontaneous oscillation.

CHAPTER V

THE DETECTOR STAGE

THE function of the Detector Stage is to convert the modulated high-frequency E.M.F., which it receives from the preceding stage, into an alternating E.M.F. of modulation frequency. It is clear that such a change cannot be brought about by any conductor which obeys Ohm's Law, so that the principal characteristic of a detector stage is that it shall be a device in which the current flowing is not proportional to the applied potential. Perhaps the simplest arrangement which fulfils this requirement consists of a pointed metallic conductor resting lightly on the surface of a crystal of Galena, Carborundum, or certain other natural or synthetic minerals. These crystal detectors were widely used in the early days of wireless telephony, but are losing their popularity owing to the fact that they require careful adjustment of the point of contact between metal and crystal; furthermore, this adjustment may be completely upset by the slightest vibration. For this reason crystal detectors are seldom

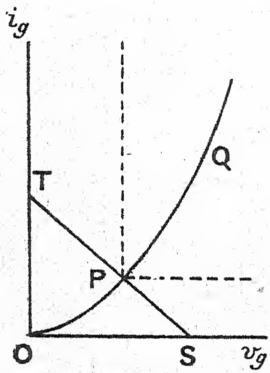

FIG. 25.

incorporated in any but the most simple receivers, and we shall not here consider them further. For a more detailed account of these detectors the reader is referred to a series of articles by Colebrook (16).

Turning next to the ordinary thermionic triode, we see that several portions of its characteristic curves are non-linear, and so might conceivably be used in a detector stage. Neglecting curvature due to saturation of the electron current, which could not be used without damage to the valve except with pure tungsten filaments, the number of possibilities reduces to three, viz.: the initial curvature of the grid current–grid voltage, anode current–grid voltage or anode current–anode voltage curve respectively The second and third of these curves are illustrated in Fig. 2 (*a*) and (*b*) respectively (p. 9), while the first is of the form indicated by the curve OPQ of Fig. 25.

Grid Circuit Detection

The process which makes use of this grid current–grid voltage curve in a detector stage is usually referred to as "Cumulative Grid" or "Leaky Grid" detection. Both of these terms are misnomers arising from an imperfect understanding of the action of the circuit, so we shall here refer to the method as *Grid Circuit Detection*. Two variations of the circuit used with this method of detection are illustrated in Fig. 26 (*a*) and (*b*) respectively. Since the principles involved are practically the same in both cases, we shall consider only the circuit of Fig. 26 (*a*). Here the H.F. input potential is applied between A and B; C is a capacity and R a resistance, the values of which we shall consider later, C_1 the input capacity, and R_1 the input resistance of the valve.

The general theory of grid circuit detection has been given by several writers, amongst whom may be mentioned Appleton and Taylor (17), Chaffee and Browning (18), Colebrook (19), and Ballantine (20). Considerable

complication arises from the fact that the grid current-grid voltage curve cannot be represented by any simple mathematical expression. Different writers make different assumptions as to the form of this curve, but in most cases the resulting theory applies only when the input E.M.F. is very small. On account of these limitations we shall here consider only a very simple case which reduces the mathematical complexities to a minimum, but which shows clearly the functions of the various portions of the circuit, and gives as much information with regard to the optimum conditions of

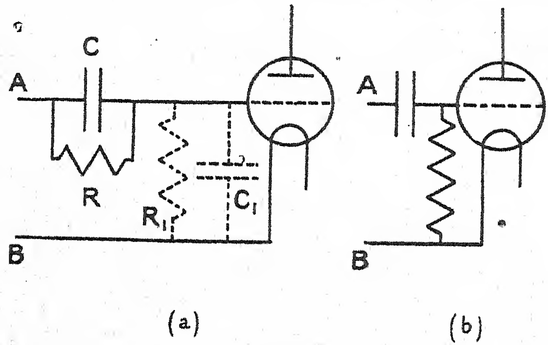

FIG. 26.

operation as can be obtained from the more detailed theories.

It will be assumed that the input E.M.F. e, which is applied between A and B (Fig. 26 (a)), is extremely small—say less than 0·1 volt—and furthermore, that the output circuit of the preceding stage, which is connected between A and B, is such as to offer negligible resistance to the flow of direct or low-frequency alternating current. (The non-fulfilment of this latter condition would entail only a very slight modification of the theory.) Since the detector action depends upon the flow of grid current, the grid bias battery will be such as to make the point

B positive with respect to the filament. With the assumptions made above, the point A will be at the same potential as B, but the grid of the valve will be at some lower value owing to the fall of potential caused by the flow of grid current through R. This initial steady potential of the grid is best determined by a graphical method in the following manner. In Fig. 25, on the axis of v_g, mark off OS to represent the steady potential v_b of the points A and B. On the axis of i_g let OT represent the current v_b/R, then the straight line TS is the graph of the equation

$$i_g = (v_b - v_g)/R,$$

and the point P, where this line cuts the grid current–grid voltage curve will indicate the steady potential of the grid, and the steady grid current flowing when no E.M.F. is applied between A and B.

Taking the simplest possible case, let us now suppose that a small high-frequency sinusoidal E.M.F., $e = E \sin \omega t$, is applied between A and B. Furthermore, let us assume that the relevant portion of the grid current–grid voltage curve in the neighbourhood of the point P can be represented by a second degree equation. It is to be noted that this is not at all the same thing as assuming that the whole curve can be represented by such an equation. It will simplify the algebra if we take P to be our new origin with axes parallel to the previous axes. The equation of the curve may then be written

$$i_g = av_g + bv_g^2 \qquad . \quad . \quad . \quad (25)$$

It is first of all necessary to calculate the high-frequency E.M.F. between grid and filament of the valve which results from the application of the input E.M.F. e, and for this purpose we must obtain an expression for the slope resistance R_g, due to electron conduction, in the neighbourhood of the working-point P (Fig. 25). By differentiation of (25),

$$\frac{1}{R_g} = \frac{di_g}{dv_g} = a + 2bv_g.$$

For very small signals we may neglect the term in v_g and write

$$R_g = 1/a.$$

Now R_g will be in parallel with R_1 and C_1 (Fig. 26), and on substitution of probable values for these terms it will be found that the impedance of the combination is considerably greater than the reactance of C_1 provided the capacity of the latter be not less than 0·0001 μf. Under this condition then, it will be sufficiently accurate if we assume the whole of the input high-frequency E.M.F. to be applied between grid and filament of the valve.

This E.M.F. will cause the potential of the grid to fluctuate by equal amounts above and below the steady value represented by P (Fig. 25), but owing to the non-linearity of the curve, the increase of the current above P will be greater than the decrease of current below P. This indicates that rectification will take place, and that the current due to the applied sinusoidal E.M.F. will consist of a D.C. component, together with an alternating component of the same frequency as the E.M.F., and possibly harmonics of this frequency also. Now we will suppose that the path via the condenser C and circuit connected between A and B offers negligible impedance to the flow of the H.F. components of the current, so that these will have no effect on the potential of the grid. The D.C. component, on the other hand, will flow through the resistance R and thus cause a decrease v_o, in the mean grid potential. Let i_o be the value of the D.C. component, then $v_o = Ri_o$, and the total E.M.F. applied to the grid as a result of the input $e = E \sin \omega t$ is equal to $E \sin \omega t - Ri_o$. Substituting in (25)

$$i_g = a(E \sin \omega t - Ri_o) + b(E \sin \omega t - Ri_o)^2.$$

Equating the D.C. components on the two sides of the equation, and remembering that $\sin^2 \omega t = \frac{1}{2}(1 - \cos 2\omega t)$

$$i_o = -aRi_o + bR^2 i_o^2 + \tfrac{1}{2}bE^2,$$

$$bR^2 i_o^2 - (1 + aR)i_o + \tfrac{1}{2}bE^2 = 0,$$

and

$$i_o = [(1 + aR) \pm \sqrt{(1 + aR)^2 - 2E^2b^2R^2}]/2bR^2.$$

Since E is small, we expand the quantity under the radical and neglect all terms except the first two. Then

$$i_o = [(1 + aR) \pm (1 + aR)\{1 - E^2b^2R^2/(1 + aR)^2\}]/2bR^2.$$

In making the choice of sign we note that the positive sign would lead to a finite value of i_o when $E = 0$. This is obviously incorrect, so we choose the negative sign giving

$$i_o = \tfrac{1}{2}E^2b/(1 + aR).$$

Remembering that $a = 1/R_g$, we may write

$$i_o = \tfrac{1}{2}R_gbE^2/(R + R_g) \quad . \quad . \quad . \quad (26)$$

From the form of this equation, we may consider the rectified current i_o as being due to a fictitious E.M.F. of magnitude $\frac{1}{2}R_gbE^2$ generated by an appliance of internal resistance R_g acting in a circuit of external resistance R.

Let us now consider the case where the input to the detector stage consists of a modulated sinusoidal E.M.F. of the form

$$e = E(1 + m \sin pt) \sin \omega t.$$

The analysis for this case could be carried out in exactly the same manner as for a sinusoidal E.M.F. ; but, since the rectified current will now contain low-frequency alternating components as well as a direct component, it can be represented as being due to not one but a series of fictitious generators, each generator being responsible for one component. Each generator will have an internal resistance R_g and, provided only a negligible fraction of the total L.F. current pass through the grid condenser C (Fig. 26), each will work into an external resistance R. (At low frequencies the input impedance of the valve may be considered infinite for our present purpose.)

In order to calculate the E.M.Fs. of these various generators, we note from equation (26) that, for the D.C. component, the E.M.F. may be obtained by multiplying R_g by the D.C. current which flows when R is equal to zero. It may easily be shown that this holds also for the other components, so we need only consider the comparatively simple case for which $R = 0$. We then have for the total grid current,

$$\begin{aligned} i &= aE(1 + m\sin pt)\sin\omega t + bE^2(1 + m\sin pt)^2\sin^2\omega t \\ &= aE(1 + m\sin pt)\sin\omega t \\ &\quad + \tfrac{1}{2}bE^2(1 - \cos 2\omega t)\left[1 + 2m\sin pt + \frac{m^2}{2}(1 - \cos 2pt)\right]. \end{aligned}$$

From this equation it is clear that the grid current contains numerous H.F. and L.F. components; for our present purpose the following are of chief interest :—

(*a*) A D.C. component of magnitude

$$\tfrac{1}{2}bE^2(1 + m^2/2).$$

(*b*) A component of frequency $p/2\pi$, the amplitude of which is bmE^2.

(*c*) A component of frequency p/π, the amplitude of which is $\tfrac{1}{4}m^2bE^2$.

The first of these is important only in that it changes the effective value of the steady grid potential; the second is the useful L.F. signal current, and we note that its magnitude is proportional to the square of the amplitude E of the carrier wave; the third is a harmonic of the useful L.F. current and is undesirable, since it represents distortion of the wave-form of the latter. If we take the ratio of harmonic to fundamental as a measure of distortion, we see that this ratio is equal to $m/4$ and thus depends only on the percentage of modulation.

In accordance with the theory given above, the magnitude of the useful L.F. current when the resistance R (Fig. 26) is not equal to zero, will be $R_g bmE^2/(R+R_g)$

and therefore the useful low-frequency E.M.F. developed between grid and filament of the valve will be

$$RR_g bmE^2/(R+R_g) \quad . \quad . \quad . \quad (27)$$

it being assumed that the circuit between A and B offers negligible impedance to the flow of L.F. current.

The functions of the circuit elements of Fig. 26 are now clear. The purpose of the condenser C is to provide a free passage for the flow of H.F. current so that practically the whole of the input E.M.F. shall be applied between grid and filament of the valve. To this end the reactance of C must be small compared with the total impedance, consisting of the input impedance of the valve in parallel with the slope grid resistance R_g. The purpose of the resistance R is obvious from equation (27), since without it no useful L.F. potential would be developed between grid and filament. Furthermore, equation (27) indicates that this potential will increase as R increases, and will reach the limiting value $R_g bmE^2$, when R becomes indefinitely great. This conclusion, however, depends upon the assumption that none of the L.F. current passes through the condenser C, and this will not be true if the product RC be too large. To allow for this shunting effect of the condenser, we must substitute for R in equation (27) the expression $R(1-jRCp)/(1+R^2C^2p^2)$, which represents the impedance at frequency $p/2\pi$ of C in parallel with R. The chief importance of this effect lies in the fact that its magnitude varies with the modulation frequency, and therefore, if it be allowed to become appreciable, it will introduce frequency distortion by causing a greater reduction in the amplitudes of the high audio frequencies than of the low. Substitution of practical values in the above expressions shows that this distortion will be unimportant if C and R be not greater than 0·0001 μf and 0·25 megohm respectively. Also, with this value of C practically the whole of the input high-frequency E.M.F. will be applied between grid and filament of the valve.

From the foregoing outline of grid circuit rectification, it will be clear that a modulated H.F. input E.M.F. gives rise to sundry high- and low-frequency E.M.Fs. developed between grid and filament of the detector valve. Now an E.M.F. v_g between grid and filament is equivalent to an E.M.F. μv_g in the anode circuit of a valve, and this will hold true for each of the E.M.Fs. under consideration, so that currents of various frequencies will flow in the anode circuit of the detector valve. If a suitable impedance be placed in the anode circuit, E.M.Fs. corresponding to each of these currents will be set up between the ends of the impedance, and may be applied to the following L.F. amplifier stage. Now the H.F. components of E.M.F. are not wanted, and it is very desirable that they should be prevented from entering the L.F. stage where, owing to coupling through stray capacities, they may give rise to spontaneous oscillation or cause distortion. The required result may be attained by using in the anode circuit a device which has a low impedance for H.F. currents and a high impedance for L.F. currents. Another advantage resulting from the provision of a low impedance path for the H.F. components of anode current is that the input impedance of the detector valve to H.F. currents will thereby be kept as high as possible (cf. p. 18).

Anode Bend Detection

Before discussing the merits or otherwise of grid circuit rectification, it will be convenient to consider the second important method of detection which is usually known as *Anode Bend Detection*. This method makes use of the curvature of the anode current–grid voltage characteristic. A curve of this type is shown in Fig. 27 (*a*), the detector circuit in Fig. 27 (*b*) and the equivalent circuit in Fig. 27 (*c*). The E.M.F. of the grid bias battery is such as to bring the steady grid potential to some point such as P on the bend of the curve (Fig. 27 (*a*)). It is here assumed that the anode circuit impedance consists

of a resistance R shunted by a capacity C: the magnitude of C is such that it has a low reactance for H.F. currents, while its shunting effect is negligible at audio frequencies. It will be assumed that C and R include the input capacity and resistance, respectively, of the following L.F. stage, with which they are in parallel. It is clear that the circuit here considered is identical with the input circuit used in grid circuit rectification, and this identity is not greatly affected if other anode circuit impedances be used in the case of anode bend detection. The essential difference between the two methods may be expressed by saying that in a grid

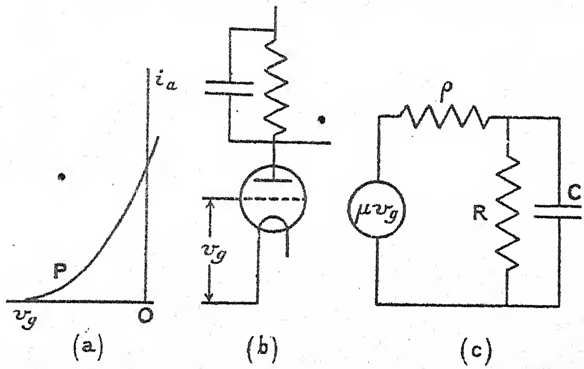

FIG. 27.

circuit detector the signal is amplified after rectification, while in an anode bend detector the amplification takes place before rectification.

COMPARISON OF DETECTORS: POWER DETECTORS

Although both of the above methods of detection have been known almost since the introduction of thermionic valves, opinion has varied considerably as to the relative merits of the two. In the early days of wireless when transmitting stations were not very powerful, and

receiving apparatus was inefficient, the only criterion by which a detector was judged was its sensitivity. From this point of view, the grid circuit detector was undoubtedly the better of the two, owing to the greater curvature of the grid current–grid voltage characteristic as compared with the anode current–grid voltage curve. Furthermore, in order to obtain the greatest possible sensitivity grid condensers of about ·0003 μf capacity and resistances of the order of two megohms were generally used. With such values the shunting effect of the condenser at the higher audio frequencies was very marked, but the resulting frequency distortion was unimportant, since it was masked by the much greater distortion introduced at other stages of these early receivers. As time went on and the quality of reception improved, grid circuit detection lost favour, since it was reputed to introduce serious frequency distortion. During the past few years both types of detector have been carefully studied, and as a result two facts have emerged : first, that with suitable choice of values of components neither type will introduce serious frequency distortion, and second, that the amplitude distortion which either introduces is far more important than the frequency distortion. At this point it will be convenient to define a measure of amplitude distortion. Suppose that owing to distortion, the current or voltage output from a stage instead of being sinusoidal, is composed of a fundamental term of amplitude E_1 and harmonics of amplitudes E_2, E_3, E_4 . . . respectively. Then we take as a measure of the distortion the ratio

$$\sqrt{(E_2{}^2 + E_3{}^2 + E_4{}^2 + \ldots)/E_1{}^2}, \qquad (28)$$

which is usually expressed as a percentage. It often happens that the only two harmonics which are important are the second and third, and in these cases it is usual to express the distortion in terms of the percentage ratios which the amplitudes of these two harmonics considered separately bear to the amplitude of the fundamental. Now, in the theory previously given for small

signals and a parabolic characteristic it was shown that the ratio of second harmonic to fundamental is $m/4$, where m is the modulation ratio. If the signal be sufficiently small, a more detailed analysis shows that this result will be approximately true even if the characteristic be not parabolic. Thus, the only possibility of reducing distortion due to second harmonic would seem to lie in the use of larger applied signal E.M.Fs. Considerable progress in this direction has been made during the last year or two, and for full details the reader is referred to papers by Warren (21), Ballantine (22), Turner (23), and Terman and Morgan (24). The following outline will, however, indicate the principles involved.

Taking first the case of anode bend detection, it is clear that, since the anode current–grid voltage curve of a triode can be represented over an extended range of voltage only by a power series in which at least the first five terms are important, the complete analysis of any rectification problem will become hopelessly complex when large input voltages are considered. For this reason it has become customary to make use of experimental characteristics, first suggested by Appleton and Taylor (17) in which change of mean anode current is plotted against amplitude of input to the detector for an applied sinsusoidal E.M.F. It is often stated that such curves, an example of which is shown in Fig. 28, may be obtained experimentally by using an input E.M.F. of very low frequency, say 50 cycles per second, when the measurements are comparatively simple. Since, however, the shape of the curve depends upon the impedance in the anode circuit, and since this impedance varies with frequency, it is preferable to make the measurements with an input E.M.F. of the same frequency as that of the signal which is to be received.

The value of such a characteristic is obvious if we regard a modulated wave as being a sine wave of varying amplitude. Such a viewpoint is legitimate in the present instance only if the variation of amplitude takes place sufficiently slowly for the rectified current at any instant

to be given by the steady state curve of Fig. 28. This condition is always satisfied in practice by suitable choice of the magnitudes of the circuit elements, since failure in this respect would lead to frequency distortion by attenuation of the higher audio frequencies.

Returning now to Fig. 28, it will be clear that if A represent the amplitude of the applied carrier wave, while B and B′ are the limits of variation of amplitude due to modulation, then, owing to the curvature of the

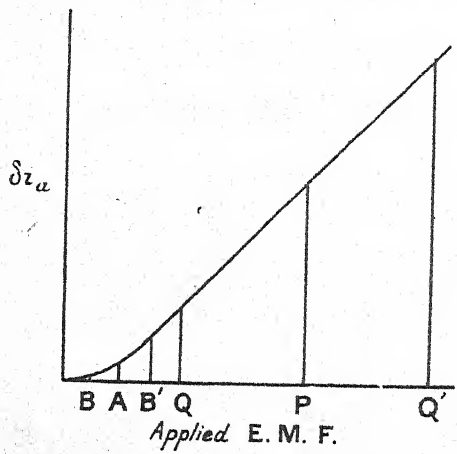

Fig. 28.

characteristic, the change of anode current will not be proportional to the change of amplitude of the applied E.M.F. If, on the other hand, the applied E.M.F. be much larger, so that P represents the carrier amplitude and Q and Q′ the limits of variation of amplitude, then since the characteristic is nearly linear between the limits Q and Q′, the change in anode current will be nearly proportional to the change of amplitude of applied E.M.F. and amplitude distortion will be much smaller than in the former case. The magnitude of amplitude

distortion can be calculated from a characteristic such as Fig. 28 by a method which will be described in Chapter VII. In general, it is found that the distortion decreases as the applied E.M.F. increases. Since a limit is set on the applied E.M.F. by the fact that the grid of the detector valve must never acquire a positive potential, a large applied E.M.F. can only be used when the steady anode potential is fairly high. With a high-tension supply of 200 volts, input E.M.Fs. of the order of 10 volts may be used when a suitable valve is chosen. Detector stages to which large input E.M.Fs. are applied are usually known as *Power Detectors*.

Turning now to power grid circuit detection, it will be clear that the general principles outlined above apply to this case also. There are, however, two additional points which must be considered. It will be remembered that, when a modulated sinusoidal E.M.F. is applied to a grid circuit detector, the resulting grid current includes a direct component. When the applied E.M.F. is large this component will also be large, and as a result the mean value of grid potential will be reduced to such an extent that grid current flows only during a small portion of each cycle. In consequence, the effective differential grid resistance will be very much larger than it was for small values of applied E.M.F., and we must take account of this fact when calculating values of grid circuit resistance and capacity. It will be found that with $C = 0{\cdot}0001\ \mu f$ and $R = 10^5$ ohms, the frequency distortion is not important.

The maximum allowable input E.M.F. in the case of power grid circuit detection is governed by the fact that the variation of grid potential must not be sufficiently large to exceed the limits of the sensibly linear portion of the anode current–grid voltage characteristic. Two undesirable effects would result from appreciable curvature of this characteristic. In the first place the low-frequency E.M.F. produced between grid and filament would be amplified in a non-linear manner and amplitude distortion would result. Secondly, the modulated

H.F. voltage which is produced between grid and filament would be rectified by the ordinary process of anode bend detection and the resulting L.F. current in the anode circuit would be out of phase with the L.F. current due to grid circuit rectification. The maximum H.F. input E.M.F. which can be applied to a power grid circuit detector varies considerably with the type of valve used, but is not usually greater than about 3 volts. In order that overloading shall not occur with voltages of this order, it is essential that the H.T. supply shall be fairly high—certainly not less than about 150 volts—whereas in the case of grid circuit detectors for very small signals it is customary to use a low value of H.T. voltage, thereby obtaining a characteristic giving maximum sensitivity.

The relative merits of the two types of power detectors described above will, of course, depend upon the characteristics of the valves used. With present-day valves the grid circuit detector seems to be the better of the two, since it is more sensitive and gives less distortion. As would be expected, indirectly heated valves make excellent detectors, since their equipotential cathodes give characteristics with sharper bends than can be obtained with directly heated valves.

Reaction

So far nothing has been said concerning the relation between the detector stage and the preceding H.F. stage. When the detector is of the anode bend type, the only action which it will have upon the H.F. stage will be to impose a slight load due to Miller effect. With grid circuit detectors, however, the input resistance is fairly low, chiefly owing to the flow of electron current between grid and filament and to the presence of the grid circuit resistance. If this input resistance should be in parallel with a tuned circuit in the preceding H.F. stage the resultant damping will make the latter very inefficient A method of overcoming this difficulty is known by the

general name of *Reaction*, and may best be explained with reference to Fig. 29. Let LRC be the tuned circuit of a H.F. stage, which is supplying the input E.M.F. to a detector valve V. Since we shall be concerned only with H.F. components of current and voltage, the grid circuit condenser and resistance in the detector stage have been omitted. Let E be the E.M.F. acting in, and i_1 the current flowing round the tuned circuit. Then the H.F. voltage applied between grid and filament will be $i_1/j\omega C$. If we assume that the reactance to H.F. currents of the anode circuit impedance is small compared with the slope resistance of the valve, we may write for the H.F. current, i_2, in the anode circuit

$$\mu i_1/j\omega C\rho \qquad (29)$$

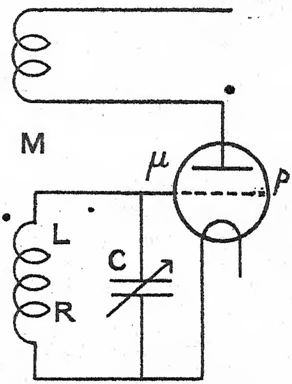

FIG. 29.

Let this current pass through a coil which is coupled to the coil of the tuned circuit to form a mutual inductance M. Then neglecting for the moment the damping effect of the valve on the tuned circuit, we may write

$$i_1[R + j(L\omega - 1/C\omega)] + jM\omega i_2 = E.$$

Substituting from (29)

$$i_1[(R + M\mu/C\rho) + j(L\omega - 1/C\omega)] = E.$$

If the input circuit be tuned to resonance so that $LC\omega^2 = 1$, then

$$i_1 = E/(R + M\mu/C\rho) \quad . \quad . \quad . \quad (30)$$

Now, if the direction of winding of the coils be such that M is negative, the effect of the anode circuit reaction coil is to reduce the resistance of the tuned circuit by an amount $M\mu/\rho C$. It is clear that, by suitable choice

of values this reduction can be made to compensate for the effective addition of resistance to the tuned circuit due to the damping caused by the detector stage. As the magnitude of M is increased, the effective resistance of the tuned circuit will decrease and eventually will become zero when instability will result. Even before spontaneous oscillation sets in the resonance curve will become so sharp that considerable frequency distortion will be caused by side-band cut-off. The principle of reaction is a very general one, and may be applied to H.F. amplifiers as well as to detector stages; furthermore, the coupling between an anode circuit and the preceding grid circuit may be either electromagnetic or electrostatic. An example of the latter type of coupling is to be found in the Miller effect with inductive anode circuit impedances. As we have seen, the effect in this case is sufficient to cause instability unless special precautions be taken, so that it is rarely necessary to provide for any additional reaction in H.F. amplifiers.

Although the circuit of Fig. 29 (*a*) is very simple, it possesses certain practical disadvantages. For example, the provision of two coils which can be moved relatively to each other necessitates a somewhat clumsy mechanical construction and makes the control of reaction not very smooth. Various alternative methods of controlling reaction have therefore been devised, and for details of these the reader is referred to the bibliography (25).

CHAPTER VI

LOW-FREQUENCY AMPLIFICATION

IF a low-frequency E.M.F. be applied between grid and filament of a triode, then a corresponding alternating voltage will be developed between the terminals of any impedance placed in the anode circuit of the valve. Since all L.F. amplifiers depend upon this principle, it is clear that they bear a formal resemblance to the H.F. amplifiers described in Chapter IV. Apart, however, from the obvious changes necessitated by the use of very much lower frequencies in the present case, there are two important differences between the two types of amplifiers. The first of these arises from the fact that whereas in the case of H.F. amplifiers the limiting values of the range of frequencies over which uniform amplification is required for any given signal differ from the mean frequency by only 1 or 2 per cent. ; in the case of audio frequency amplifiers it is essential that the amplification should be as uniform as possible over a range of frequency extending from 30 to 10,000 cycles per second. The use of any tuned circuit as anode impedance is therefore out of the question in the latter case. The second difference will be apparent from the following considerations. Since a H.F. amplifier is followed by a detector stage the output E.M.F. which it is called upon to deliver will rarely exceed 10 volts, and will usually be much less. An L.F. amplifier, on the other hand, is followed by a power stage, to which it may be required to deliver an E.M.F. as high as 100 volts or

more in the case of a large receiver. When dealing with H.F. amplifiers, we assumed the anode current–grid voltage characteristics of the valve to be a series of parallel straight lines, and this was legitimate over the small range of variation of grid potential there considered. When dealing with L.F. amplifiers, however, it is necessary to enquire whether amplitude distortion is not being produced by curvature of the valve characteristics. Should such distortion become excessive the stage is said to be overloaded. The subject of overloading will be considered in detail in Chapter VII.

Resistance Capacity Coupling

There are three types of L.F. amplifier, in which the coupling from one stage to the next is effected by

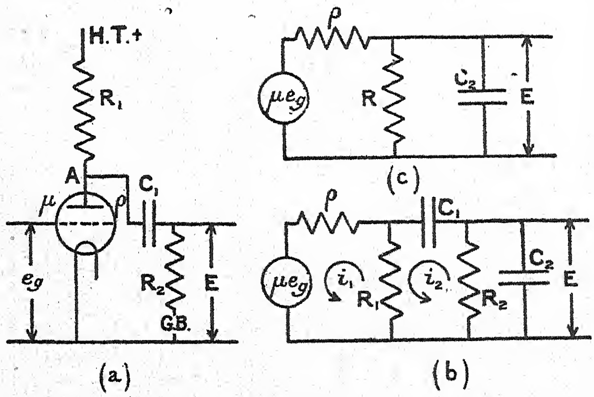

Fig. 30.

Resistance-Capacity, Choke-Capacity, and Transformer, respectively. The circuit for resistance-capacity coupling is shown in Fig. 30 (a), where an input E.M.F. e_g is applied between grid and filament of a valve of amplification factor μ and slope resistance ρ. The condenser C_1 ensures that the steady potential of the point A shall

not be applied to the grid of the following valve, and the resistance R_2 and battery G.B. ensure that the grid of this valve is correctly biassed. Obviously, the insulation of C_1 needs to be very good. In drawing the equivalent theoretical circuit, Fig. 30 (*b*), it is necessary to take account of the input impedance of the following stage; this may be represented by a condenser C_2, the capacity of which is usually of the order of 100 $\mu\mu f$, so that if R_2 be of the order of 1 megohm the shunting effect of C_2 will only be appreciable at high audio frequencies. Neglecting C_2 for the moment, the circuit equations become

$$\rho i_1 + R_1(i_1 - i_2) = \mu e_g$$
$$R_1(i_2 - i_1) + R_2 i_2 - j i_2/C_1\omega = 0,$$

where ω is the angular frequency of e_g. After suitable reduction, this becomes

$$i_2\left[R_2 + \frac{\rho R_1}{\rho + R_1} - \frac{j}{C_1\omega}\right] = \frac{\mu R_1 e_g}{R_1 + \rho} \qquad (31)$$

Since the output voltage E is equal to $R_2 i_2$, the absolute magnitude of E will be independent of frequency provided $1/C_1^2\omega^2$ be negligible compared with

$$[R_2 + \rho R_1/(\rho + R_1)]^2.$$

When this condition is not fulfilled the variation of amplification with frequency may be calculated from (31). When, however, it is fulfilled, we have

$$\text{Amplification} = \frac{E}{e_g} = \frac{\mu R}{\rho + R} \quad . \quad . \quad (32)$$

where R is equivalent resistance of R_1 and R_2 in parallel. At high audio frequencies the effect of C_1 will be negligible, but allowance must be made for the shunting effect of C_2. The equivalent circuit will now be as shown in Fig. 30 (*c*), and from this it is easy to show that

$$\mu e_g/E = j\rho C_2\omega + (\rho + R)/R$$

and therefore

$$|\mu e_g/E| = \sqrt{\rho^2 C_2^2\omega^2 + (\rho + R)^2/R^2}. \quad . \quad (33)$$

Now R will usually be considerably larger than ρ, so that $(\rho + R)/R$ will be approximately equal to unity. Thus, if the amplification is to be independent of frequency, $\rho^2C_2{}^2\omega^2$ must be small compared with unity. If we arbitrarily put it equal to 0·05 when $\omega = 5 \times 10^4$, and take $C_2 = 100$ $\mu\mu f$, we find ρ approximately equal to 50,000 ohms. Clearly ρ must not greatly exceed this value if frequency distortion is to be unimportant. In order to obtain as large an amplification as possible, μ should be made as high as the above limitation of ρ will permit.

From equation (32) we see that the amplification will only approach the limiting value provided R be large compared with ρ, and this will necessitate both R_1 and R_2 being large compared with ρ. Now it is not desirable that either R_1 or R_2 should be increased indefinitely for the following reasons. It has been tacitly assumed that no grid current flows in the valve following the stage under consideration and in general this will be true. During the reception of exceptionally loud passages of music, however, occasions may occur when the instantaneous value of the grid potential of the following valve becomes positive. On such an occasion a pulse of grid current would flow through R_2 and charge up the condenser C_1 in such a manner as to make the grid bias applied to the following valve considerably greater than the normal value due to the battery G.B. After a short interval of time all the charge given to C_1 will have leaked away through R_2 and the effective bias will have returned to its normal value. Since incorrect biasing of the following valve may cause considerable amplitude distortion, it is important that any charge given to C_1 by flow of grid current should leak away as quickly as possible. For this reason the value of R_2 is not usually made larger than about half a megohm.

Turning now to R_1, we note that the higher this resistance is made, the lower will be the potential applied to the anode of the valve, if the H.T. battery voltage be fixed. In consequence, it was formerly the practice to

limit the value of R_1 to some 100,000 ohms, since it was assumed that any larger value would reduce the anode voltage to such an extent that the straight portion of the valve characteristic would not be sufficiently long to accommodate the applied input E.M.F.; amplitude distortion would then arise owing to the curvature of the characteristic. This, however, is not the whole truth, since we are concerned not with the characteristic of the valve itself, but with the characteristic of the valve in series with R_1, and clearly the addition of R_1 will tend to straighten out any curvature in the characteristic of the valve alone. It was first pointed out by von Ardenne and Heinert (26) that the extra straightening effect obtained by increasing the resistance of R_1 compensated to a great extent for the decrease in anode voltage, and therefore that anode resistances of the order of 2 megohms could be used without serious amplitude distortion, provided the input E.M.F. were not too large. This conclusion has been entirely confirmed by a mathematical analysis due to Colebrooke (27). However, since the stage amplification depends upon the resistance of R_1 and R_2 in parallel, and since R_2 should not exceed half a megohm, there is little to be gained by making R_1 larger than this value.

Transformer Coupling

We consider next those amplifiers in which the coupling from one stage to the next is effected by means of a transformer. The actual circuit is shown in Fig. 31 (*a*) and the equivalent theoretical circuit in Fig. 31 (*b*), where L_1 is the primary inductance, r_1 the sum of the valve slope resistance and the transformer primary resistance, r_2 the secondary resistance modified to include the input resistance of the next stage, and C includes the self-capacity of the secondary winding and the input capacity of the next stage. In practice, it is not possible to obtain perfect coupling between the two windings of the transformer, so that all the magnetic lines of force

due to the primary do not pass through the secondary and vice versa. This fact may be represented, as in Fig. 31 (*b*), by supposing the secondary inductance to consist of two parts, viz. L_2, which is perfectly coupled to L_1, and l, which is not at all coupled to L_1. The total secondary inductance will then be $L_2 + l$; l is usually known as the leakage inductance. Now, since perfect coupling exists between L_1 and L_2, these inductances form an ideal transformer of ratio $s = \sqrt{L_2/L_1}$, and we may apply to them the theory developed on page 31. By this means the problem of calculating the amplification of the stage (i.e. the ratio v_2/v_g) may be divided into

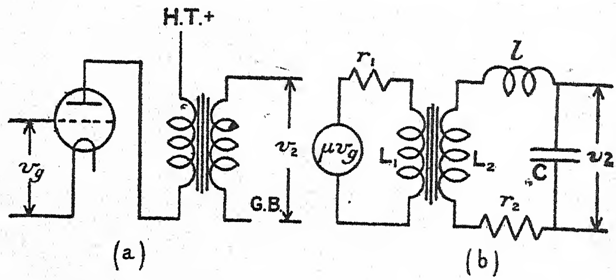

FIG. 31.

two parts. In the first place, from a consideration of the circuit of Fig. 32 (*a*) in which the secondary load has been transferred to the primary, we calculate the E.M.F. v_1 produced between the ends of the inductance L_1. Then the voltage induced between the ends of L_2 will be sv_1 and from a consideration of the circuit of Fig. 32 (*b*), the output voltage v_2 may be determined. Even with these simplifications the algebra becomes somewhat complicated, and it is difficult to estimate the relative importance of the various terms. To overcome this difficulty, it will be convenient to assign definite values to the circuit elements, and then carry out the calculations

for certain selected frequencies. As typical values we may take

$$L_1 = 50 \text{ Henries}, \; L_2 = 450 \text{ Henries}, \; l = 9 \text{ Henries},$$

$$r_1 = 10{,}000 \text{ ohms}, \; r_2 = 60{,}000 \text{ ohms}, \; C = 100 \; \mu\mu f, \text{ and } s = 3.$$

Let Z be the impedance of l/s^2, r_2/s^2 and s^2C in series: then

$$Z = 6667 + j(\omega - 10^{10}/9\omega) \quad . \quad . \quad (34)$$

At a frequency of 200 cycles per second Z will be approximately equivalent to a condenser with reactance of 1 megohm, while the reactance of L_1 will be about 60,000 ohms. At this frequency, therefore, the shunting effect

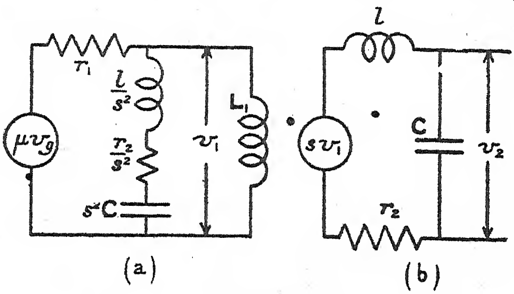

FIG. 32.

of Z may be neglected, and at lower frequencies its effect will be still less important. Thus, for all frequencies up to 200 cycles per second, v_1 may be calculated from the simple formula

$$v_1 = jL_1\omega\mu v_g/(r_1 + j\omega L_1)$$

or

$$|v_1| = \omega L_1 \mu v_g / \sqrt{r_1^2 + \omega^2 L_1^2} \quad . \quad . \quad (35)$$

At 200 cycles per second v_1 will be very nearly equal in magnitude to μv_g. As the frequency increases the reactance of L_1 will increase, while Z will decrease, but so long as both are very large compared with r_1, v_1 will not differ sensibly from μv_g, and this condition will hold up to a frequency of about 3500 cycles per second. It is

to be noted that at a frequency of about 700 cycles per second a condition of resonance exists, and the circuit consisting of Z and L_1 in parallel will then behave as a pure resistance of about 1 megohm. At this point v_1 will still be very nearly equal to μv_g, so the resonance causes no sudden change in the value of v_1. Above 3500 cycles per second, the reactance of L_1 will be very large compared with Z, and so may be neglected in calculating v_1. We then have

$$v_1 = Z/(Z + r_1) \quad . \quad . \quad . \quad (36)$$

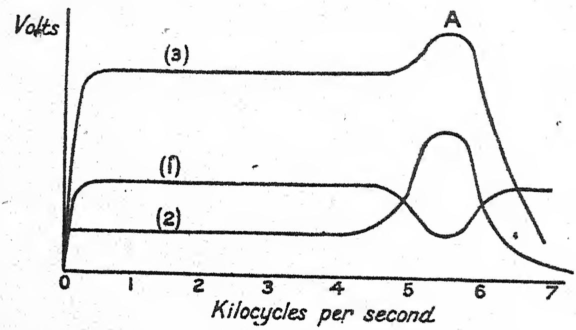

FIG. 33.

so that v_1 will be least when Z is least. Reference to equation (34) shows that a minimum value of Z will occur when the frequency is such as to cause resonance between l/s^2 and s^2C. This will occur at about 5500 cycles per second, and Z will then be equivalent to a pure resistance of 6667 ohms, giving $v_1 = 0{\cdot}4\ \mu v_g$. For still higher frequencies Z will increase again and v_1 will rise almost to the value μv_g.

From the above considerations it will be clear that if v_g remain constant, the relation between v_1 and frequency will be somewhat as shown by curve (1) of Fig. 33.

Turning now to the second part of the problem, we will first suppose that the input E.M.F. in Fig. 32 (*b*)

instead of being sv_1 is constant and equal to s volts. Then, since we are dealing with a simple series tuned circuit, the output voltage v_2' will be given by

$$|v_2'| = s/C\omega\sqrt{r_2^2 + (l\omega - 1/C\omega)^2} \quad . \quad (37)$$

and, as in the primary circuit, the resonance will occur at about 5500 cycles per second. The variation of v_2' with frequency is shown by curve (2) in Fig. 33. Finally, when the input to the circuit of Fig. 32 (*b*) is Sv_1, the output E.M.F. will be given by $v_2 = v_1v_2'$, and can be obtained by multiplying the corresponding ordinates of curves (1) and (2) in Fig. 33. The resulting variation of v_2 with frequency is shown in curve (3) of Fig. 33. It should be noted that these curves are not drawn to scale. Since v_g is assumed to be constant, curve (3) represents the variation with frequency of the overall amplification of the stage, and we must now consider what steps should be taken to keep this amplification constant over as large a range of frequency as possible. The most important conditions may be considered under three headings, as follows :—

(*a*) In order to avoid a falling-off in amplification at low frequencies the primary inductance must be large. Equation (35) shows that the falling-off will be unimportant provided the reactance of the primary be large compared with the slope resistance of the preceding valve, since in practice this slope resistance will not differ greatly from r_1.

(*b*) It is clear from curve (3) of Fig. 33 that the amplification decreases rapidly for frequencies greater than that at which the resonance peak A occurs. Therefore, by reduction of the capacity C and the leakage inductance l this resonance should be made to occur at as high a frequency as possible. The reduction of l is effected by making the coupling between primary and secondary as perfect as possible.

(*c*) In order to avoid excessively large amplification at the resonance peak A, Fig. 33, the secondary resistance may be increased (*vide* equation (37)) by the use of wire of low conductivity. Alternatively a suitable high resistance is sometimes placed in parallel with the secondary winding.

It is difficult to satisfy all these conditions and, at the same time, obtain a high overall stage amplification. Since the overall amplification is equal to μs, it is desirable that both μ and s be as high as possible. But a high value of μ involves a high slope resistance of the valve, while a high turns ratio s involves either a large secondary inductance or a small primary inductance. If the primary inductance be small, condition (*a*) above will not be satisfied, while if the secondary be large, its self-capacity will be great and it will be impossible to keep the secondary circuit capacity C within the limits prescribed by condition (*c*) above. It is thus clear that a compromise must be adopted, and with a transformer of moderate size it is found possible to obtain an overall stage amplification of about 40, while maintaining a reasonably good frequency characteristic. To obtain a high primary inductance and also to make the coupling between primary and secondary as perfect as possible, the coils are wound upon an iron core which is laminated to reduce the effects of eddy currents. Since the transformer primary is connected in the anode circuit of a valve, the current traversing it contains a direct as well

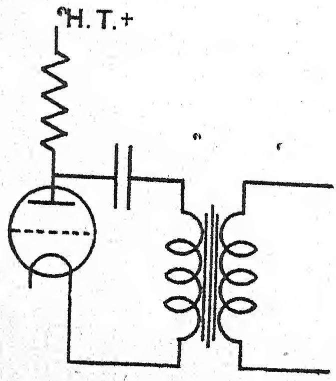

FIG. 34.

as an alternating component. This direct current produces a steady magnetic field which, unfortunately, reduces very considerably the effective permeability of the iron core, and thus decreases the value of the primary inductance. This difficulty can be avoided to some extent by using for the transformer core special alloys which are not so adversely affected by a steady magnetic field, but it is frequently advantageous to employ some such circuit as is shown in Fig. 34, where an alternative path is provided and the direct current is prevented from flowing through the primary winding.

A third method of coupling which is sometimes employed in L.F. amplifiers is known as *Choke Coupling*, and differs from Resistance Coupling only in that a choke is used in place of the anode resistance. This choke, which is wound upon an iron core, is designed to have as high an inductance and as low a self-capacity as possible.

Comparison of Types of Coupling

The choice of coupling for a particular L.F. amplifier will depend upon whether stage amplification or absence of distortion is the more important. The frequency characteristics of the best transformer coupled stages available at the present time are not as good as those which can be obtained with well-designed resistance coupled amplifiers. On the other hand, a somewhat greater stage amplification can be obtained by the use of a transformer. Choke coupling gives a frequency characteristic intermediate between the other two and, in general, has little to recommend it.

We must next consider whether the methods of coupling described above will be satisfactory when they directly follow a detector stage. In this case a low impedance path must be provided for the components of H.F. current present in the anode circuit of a detector. This is usually accomplished by connecting the anode of the detector to the filament via a small condenser of

about 0·0001 μf capacity. When reaction is used the H.F. current will, of course, traverse the reaction coil in passing from anode to filament. The use of a by-pass condenser in this manner is not entirely satisfactory since, if the capacity be made large enough to avoid the application of any appreciable H.F. voltage to the next stage, it will usually cause a certain amount of cut-off of the higher audio frequencies. This difficulty can be overcome by replacing the condenser by a properly designed high-pass filter.

When the detector stage employs grid circuit rectification either resistance coupling or transformer coupling may be used, but when anode bend rectification is employed the valve will be operating on the lower bend of its anode current–grid voltage characteristic and its slope resistance will therefore be abnormally high so that, in this case, resistance coupling is to be preferred, since a normal intervalve transformer would not have a sufficiently large primary inductance to prevent undue cut-off of the very low audio frequencies.

Push–pull Amplification

We must now consider what is known as the *Push-pull* system of amplification. An amplifier employing this principle is illustrated in Fig. 35, where the input E.M.F. v_0 is applied to a transformer, the secondary of which is tapped at its mid point and thus gives rise to two E.M.Fs., v_1 and v_1', which are equal in magnitude but opposite in phase. It will be seen that the valves are used in pairs and, since the two valves of any pair are matched as closely as possible, and the intervalve couplings are also matched, the voltages produced between corresponding points of the two branches of the amplifier will always be equal in magnitude but opposite in phase. As shown in Fig. 35, the currents derived from the last pair of valves pass respectively through the two halves of the centre-tapped primary of a transformer, the secondary of which is connected to a loud speaker. Alternatively,

the loud speaker may be arranged to have two equal windings in opposite directions and the currents from the last pair of valves may then pass directly to it. Although only one pair of amplifying stages is shown in Fig. 35, a larger number may be used when necessary, and any of the previously described forms of coupling may be used between two consecutive pairs of valves.

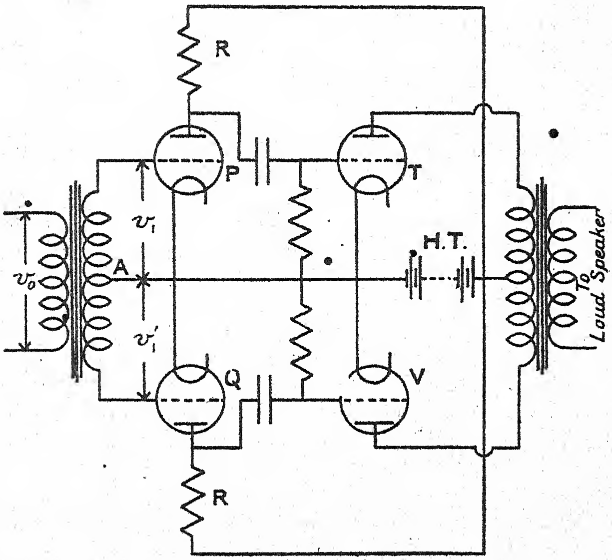

FIG. 35.

To investigate the advantages of this somewhat complicated system of amplification, we will suppose that the voltages applied to the grids of the first pair of valves in Fig. 35 are given by $v_1 = A \sin \omega t$, $v_1' = -A \sin \omega t$. We will also suppose that the characteristics of these valves are not quite linear, and that the anode current flowing through the valve and the anode circuit is related

to the grid potential by the equation $i = av + bv^2$, where a and b are constants. Then

$$i_1 = aA \sin \omega t + bA^2 \sin^2 \omega t$$
$$= aA \sin \omega t + \tfrac{1}{2}bA^2 - \tfrac{1}{2}bA^2 \cos 2\omega t$$

and $$i_1' = -aA \sin \omega t + \tfrac{1}{2}bA^2 - \tfrac{1}{2}bA^2 \cos 2\omega t.$$

Now, if the characteristics of the last pair of valves be linear, the currents flowing in their anode circuits will be proportional to i_1 and i_1' respectively, and, since the primary windings of the output transformer are arranged so that the fundamental components of the two anode currents will aid each other, it is clear that the effects due to the two second harmonic components will neutralize each other. It is easy to extend the analysis to show that, in the general case, when the curvature of the valve characteristics is such as to introduce harmonics higher than the second, then all the even harmonics are balanced out. The same principle will also hold for harmonics produced by any pair of valves other than the first, which was considered above.

From the foregoing it will be clear that the push-pull system is of most use in large amplifiers, where the voltages concerned are so large that the curvature of valve characteristics becomes important. With triodes the second harmonic is usually considerably greater than any of the higher harmonics, and its elimination is therefore particularly useful.

There are other advantages of the push-pull system which are important in the construction of large amplifiers. Thus if the device which supplies H.T. current to the receiver have an appreciable internal resistance, any variation in the current taken by one stage of a receiver will cause a corresponding fluctuation of the H.T. supply voltage, which will affect all stages of the receiver. Under certain circumstances (28) this effect may cause the receiver to become unstable. Obviously the difficulty is absent when the push-pull system is employed, since the algebraic sum of the anode currents flowing to each pair of valves is constant.

We have seen that an amplifier using resistance coupling can be designed to give a better frequency characteristic than one which employs transformer coupling, so that it is important to enquire whether a system of push-pull amplification can be arranged, from which the input transformer with tapped secondary is omitted. This has been successfully accomplished by Carpenter (29), who has conducted a detailed investigation of the subject. Of the many circuits which he has devised for the purpose, only the simplest, which is shown in Fig. 36, will be considered here. The circuit

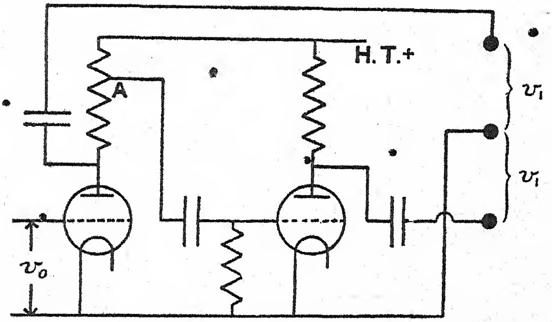

FIG. 36.

is self-explanatory, and depends upon the fact that the output and input E.M.Fs. from a single valve resistance-coupled amplifier are 180° out of phase with each other. Bearing this fact in mind, it is clearly possible, by suitable choice of the point A at which the anode circuit resistance of the first valve is tapped, to arrange that the voltages v_1 and v_1' are equal in magnitude but opposite in phase. For details of the more complicated circuits, which have several advantages over the simple one described above, the reader is referred to the original paper.

CHAPTER VII

THE POWER STAGE

WHEN the output E.M.F. from the L.F. amplifier of a receiver is applied between grid and filament of the power valve, which for the moment we will suppose to be a triode, electrical power is delivered to the loud speaker which is connected in the anode circuit of this valve. Now for a given valve and given input E.M.F., the magnitude of the power will depend upon the impedance of the loud speaker, and will therefore vary with frequency. Whether this variation is desirable or otherwise, from the point of view of quality of reproduction of sound, will depend upon the way in which the loud speaker impedance and its acoustic output vary with frequency. Since the variations of both of these quantities with frequency are usually very complicated and, moreover, differ widely from one loud speaker to another, it is clear that the problem of elimination of frequency distortion of the acoustic output can only be tackled when the characteristics of the loud speaker are known. A discussion of this problem is outside the scope of the present volume; the reader who is interested should consult a recent paper by Sowerby (30).

In the design of the power stage, therefore, the only type of distortion with which we shall be concerned is amplitude distortion arising from non-linearity of the valve characteristics and the magnitude of this distortion will be expressed by the percentage of harmonics introduced by the valve (see p. 62).

Dynamic Characteristics

In the investigation of the amplitude distortion there are three possible methods of attack. The first of these is a purely experimental one, and will be described later. The second is analytical, and involves the derivation of an expression for the distortion from the equation of the static characteristics of the valve and the constants of the loud speaker. This method is greatly limited by the fact that it is impossible to express the static characteristics of a valve by any very simple equation, which will

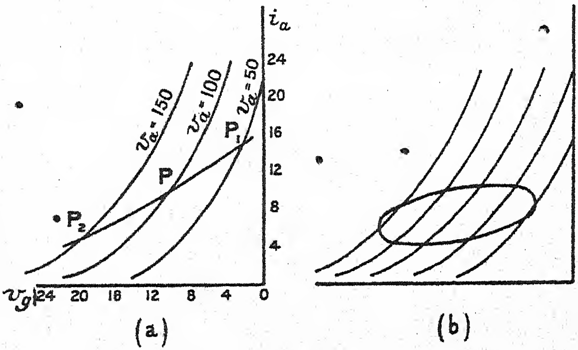

Fig. 37.

hold accurately over a sufficiently large range of values of anode current and voltage, and although it may yield useful results in certain cases, it is not of very general application and will not be considered further. We proceed to explain the third method, which is a graphical one, and involves the use of what are known as the *Dynamic Characteristics* of the valve and its associated anode circuit. The significance of this term will be appreciated from the following considerations. Suppose we are dealing with a triode, the ordinary anode current-grid voltage characteristics of which are represented to scale by the full lines in Fig. 37 (*a*). Let the anode circuit

consist of a resistance of 10,000 ohms connected between the anode of the valve and the positive terminal of the H.T. supply. Suppose the voltage of the H.T. supply to be 200 volts, and let the grid of the valve be biassed to a steady potential of $-$ 10 volts. Then, if the anode current flowing under these conditions be 0·01 ampere, the potential of the anode will be 200 $-$ (10,000 $\times$ 0·01) = 100 volts. Clearly this state of affairs may be represented by the point P in Fig. 37 (*a*). A point such as P which represents the conditions existing at any particular instant is known as an *Operating Point.* Now suppose that, owing to a change in the grid voltage, the anode current increases to 0·015 ampere. Then the anode potential will decrease to 200 $-$ (10,000 $\times$ 0·015) = 50 volts, and the operating point will move to P_1, thus showing that the new grid voltage required to effect this change is about $-$ 2 volts. Similarly, if the change in grid voltage had been such as to cause a decrease in anode current to 0·005 ampere, the operating point would have moved to P_2. A curve such as P_1PP_2, which represents the locus of the operating point when the grid voltage is varied, is known as a Dynamic Characteristic. Its position and shape depend upon the nature of the anode impedance as well as upon the properties of the valve itself. When the anode circuit impedance is reduced to zero the dynamic characteristics coincide with the ordinary static characteristics of the valve itself. In the case considered above where the anode circuit impedance is a pure resistance, it is clear that, if the valve static characteristics be equidistant parallel straight lines over the portion considered, the dynamic characteristic will also be a straight line, but when the static characteristics are not equidistant parallel straight lines, then the dynamic characteristic will be curved. In either case the dynamic characteristic corresponding to any particular value of anode circuit resistance may be plotted point by point, as shown above. When the anode circuit impedance has reactance as well as resistance, the dynamic characteristic cannot in general be plotted,

since the anode current and anode potential depend not only on the grid voltage, but also on the rate at which the grid voltage is being varied. When the grid voltage varies sinusoidally with respect to time and the static characteristics of the valve are equidistant parallel straight lines, the anode current will also be a sinusoidal function of time and there will be a constant phase difference between the anode current and the anode potential. In this special case the dynamic characteristic will be an ellipse as shown in Fig. 37 (*b*), and the eccentricity of this ellipse will depend upon the ratio of resistance to reactance in the anode circuit, and will therefore vary with frequency. In practice a loud speaker will be connected in the anode circuit of the valve and the static characteristics of the latter will not, in general, be parallel straight lines, so that the dynamic characteristics of the combination will be a series of distorted ellipses. Such curves are too complicated to admit of useful graphical treatment, so it becomes necessary to introduce some simplifying assumption. The plan which is generally adopted is to assume that the loud speaker behaves as a pure resistance, the magnitude of which is equal to the actual impedance of the loud speaker at the frequency under consideration. This assumption is justified by the fact that it leads to a method for the comparison of the distortions produced by different output stages, and this method yields results in agreement with practical experience. Now, for a given input E.M.F., the distortion introduced by the output stage will depend upon the value of the loud speaker impedance, and will therefore vary with frequency, so that in designing an output stage we must base our calculations on that frequency at which the greatest average output signal amplitudes may be expected to occur. Very little information seems to be available concerning the distribution of average amplitude with frequency in normal wireless transmissions of speech and music, but it would probably be fairly safe to design the output stage for some intermediate frequency in the

neighbourhood of 1000 cycles per second. In future, therefore, we shall suppose a loud speaker to be equivalent to a fixed pure resistance of magnitude equal to the actual impedance of the loud speaker at the somewhat arbitrarily chosen frequency of 1000 cycles per second.

We shall suppose the loud speaker to be connected to the power valve through a transformer, as shown in Fig. 38 (*a*). This arrangement isolates the loud speaker from the source of H.T. supply, and prevents the direct component of the anode current from passing through the coils of the loud speaker. When the transformer is of unit ratio it may be replaced by a choke and condenser, as shown in Fig. 38 (*b*). In general, if S be the

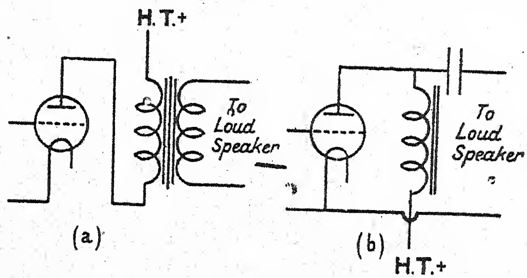

FIG. 38.

ratio of the transformer, and R the equivalent resistance of the loud speaker, the combination of transformer and resistance may be replaced, when we are considering the output stage, by a resistance of magnitude R/S^2 (see p. 31). Hence, whatever the actual value of the equivalent resistance of the loud speaker, its effective value from the point of view of the output stage may be varied at will by suitable choice of S.

It must be remembered that the replacement of loud speaker and transformer by a resistance is only valid when we are considering the flow of alternating current. So far as the flow of the direct component of anode current is concerned, the only resistance in the anode circuit is

the actual resistance of the primary of the transformer, and this is usually quite negligible in comparison with the slope resistance of the valve. Hence, when no signal E.M.F. is applied to the grid of the output valve, the potential of the anode will be equal to the voltage of the H.T. supply. On the application of a signal the potential will oscillate about this initial value and will sometimes be very much greater than the voltage of the H.T. supply.

Amplitude Distortion

A number of variable factors enter into the design of an output stage, and we must now treat these in some

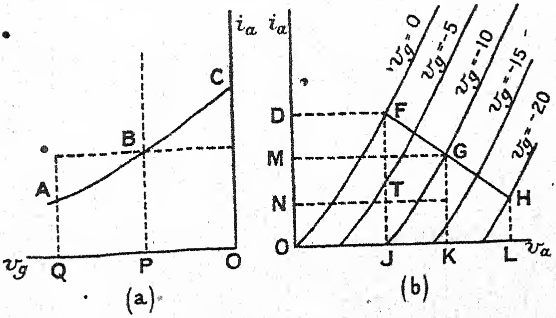

Fig. 39.

detail. In the first place, experience shows that it is always desirable for the H.T. voltage applied to the valve to be as great as possible. It will therefore be assumed in future that this voltage is fixed either by limitations of the source of supply or by the manufacturer's rating of the valve. Let us suppose that for a certain output valve with a given resistance load in its anode circuit the dynamic characteristic is as represented by curve ABC of Fig. 39 (*a*). Now, from a consideration of the method of derivation of this curve and of the shapes of the static characteristics of a triode, it will be

clear that the curvature of the dynamic characteristic tends to increase as the grid potential becomes more and more negative. To avoid excessive distortion, the curvature of the dynamic characteristic must be as small as possible, so that the grid bias potential should be fixed at a value as little negative as is consistent with the condition that signals never cause the grid to acquire a positive potential. In other words, the negative grid bias should be numerically equal to the peak voltage of the applied signal E.M.F., so that grid current does not flow at any portion of the cycle. Then, in Fig. 39, let us suppose the steady grid bias to be such that the initial operating point is B, and that a signal voltage $v = a \sin \omega t$ causes the operating point to swing between the limits A and C. Now, for a triode with resistance load working under conditions for small distortion, the dynamic characteristic may be represented with sufficient accuracy by the equation

$$i = bv + cv^2$$

where b and c are constants and the point B is taken as origin of co-ordinates. Substituting for v we find

$$i = ba \sin \omega t + \tfrac{1}{2}ca^2(1 - \cos 2\omega t) \qquad (38)$$

so that

$$\frac{\text{Amplitude of Second Harmonic}}{\text{Amplitude of Fundamental}} = \frac{ca}{2b}.$$

Our representation of the dynamic characteristic by a parabolic equation involves the assumption that harmonics higher than the second are absent, so that the above fraction is a measure of the distortion; we note that it is proportional to the applied signal E.M.F.

Returning to equation (38), when

$$\omega t = \pi/2,\ v = a \text{ and } i_{\text{max.}} = ba + ca^2.$$

Also when

$$\omega t = 3\pi/2,\ v = -a \text{ and } i_{\text{min.}} = ca^2 - ba.$$

Referring to Fig. 39 (a), and remembering that B is the origin of co-ordinates, we see that

$$\mathrm{OC} - \mathrm{BP} = ba + ca^2$$
$$\mathrm{BP} - \mathrm{AQ} = ba - ca^2.$$

Therefore,

$$\text{Distortion} = \frac{ca}{2b} = \frac{\mathrm{OC} + \mathrm{AQ} - 2\mathrm{BP}}{2(\mathrm{OC} - \mathrm{AQ})} \quad . \quad (39)$$

Also, the amplitude of the fundamental component of anode current

$$= ba = \tfrac{1}{2}(\mathrm{OC} - \mathrm{AQ}) = \tfrac{1}{2}(i_{\text{max.}} - i_{\text{min.}}).$$

Therefore the fundamental power developed in the anode load of magnitude R

$$= \frac{\mathrm{R}}{8}(i_{\text{max.}} - i_{\text{min.}})^2$$
$$= \tfrac{1}{8}(i_{\text{max.}} - i_{\text{min.}})(\mathrm{E}_{\text{max.}} - \mathrm{E}_{\text{min.}}),$$

where $\mathrm{E}_{\text{max.}}$ and $\mathrm{E}_{\text{min.}}$ are the limits of variation of anode potential.

One objection to the above method of calculating distortion is that the construction of the requisite dynamic characteristic is a somewhat tedious process. This difficulty is eliminated if we make use of a set of static characteristics in which anode current is plotted against anode potential for a number of values of grid potential. Referring to Fig. 39 (b), if K represent the steady H.T. supply voltage and the grid be biassed to a potential of $-$ 10 volts, then G will be the initial operating point. If the application of a signal E.M.F. cause the grid potential to swing between zero and $-$ 20 volts, then, since at any instant the increase in anode current multiplied by the load resistance R must be equal to the decrease in anode potential, the dynamic characteristic will be the straight line FGH where TH/TF = R. In general, FG will not be equal to GH ; let us put

FG/GH $= p$. Then, substituting the corresponding values in equation (39) we find

$$\text{Distortion} = \frac{\text{FJ} + \text{HL} - 2\text{GK}}{2(\text{FJ} - \text{HL})},$$

and by similar triangles this is equal to

$$\frac{\text{FG} - \text{GH}}{2(\text{FG} + \text{GH})} = \frac{(p-1)}{2(p+1)}.$$

Now, it is generally accepted that the output stage is satisfactory if the distortion be less than 5 per cent., and from the above equation we see that this limiting value will occur when $p = 11/9$. As before, the output power will be $\frac{1}{8}(E_{\text{max.}} - E_{\text{min.}})(i_{\text{max.}} - i_{\text{min.}})$, and this is equal to one-quarter of the area of triangle FTH.

Optimum Resistance Load.

Since the effective resistance load in the anode circuit of the output valve may be adjusted to any required value by suitable choice of the ratio of the output transformer, it is important to decide what resistance will enable the greatest output power to be obtained for a given degree of distortion. This problem was first solved by Brown (31), who assumed that the anode current-grid voltage static characteristics of a triode are equidistant parallel straight lines provided the anode current do not fall below some limiting value $i_{\text{min.}}$ Then, referring to Fig. 40, one end of the dynamic characteristic must lie on OL, which is the ordinate erected at zero grid potential, while the other lies on the line MN, below which the characteristics become curved. With the assumption made as to the shape of the characteristics, these conditions represent the limits of variation of grid potential if distortion is to be avoided. Furthermore, so long as these conditions be observed, any dynamic characteristic such as CBD will be a straight line. Let static characteristics for anode voltages $E_{\text{max.}}$ and $E_{\text{min.}}$

be drawn through C and D respectively. The positions of these points will depend upon the magnitude of the anode load. Midway between these characteristics will lie that for anode voltage E_o, which is fixed, and passes through the initial operating point B. Now the A.C. output power is proportional to

$$(E_{max.} - E_{min.})(i_{max.} - i_{min.})$$
$$\text{or to } (E_o - E_{min.}) \,.\, DN$$
$$\text{or to } GD \,.\, DN.$$

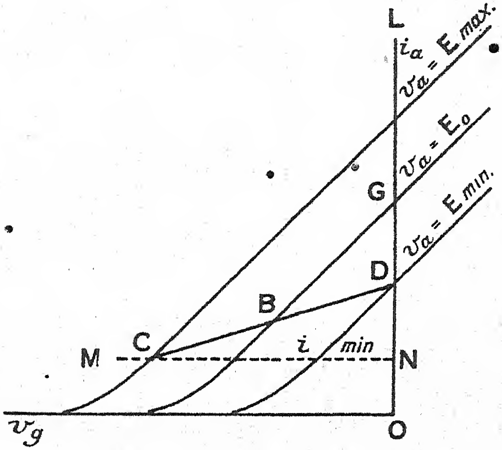

FIG. 40.

Since G and N are fixed, the product will be maximum when GD = DN. If ρ be the slope resistance of the valve,

$$GD = (E_o - E_{min.})/\rho = (E_{max.} - E_{min.})/2\rho.$$

Hence, the output will be a maximum when

$$(E_{max.} - E_{min.})/2\rho = DN = (i_{max.} - i_{min.}).$$

But R, the resistance of the anode load, is equal to

$$(E_{max.} - E_{min.})/(i_{max.} - i_{min.})$$

so that the condition becomes

$$R = 2\rho.$$

Although this result is based upon assumptions which are only approximately true, its validity has been confirmed by experimental investigations (32).

Let us suppose that an alternating E.M.F. e_g is applied between grid and filament of a triode of amplification factor μ and slope resistance ρ, which has a resistance load R connected in its anode circuit. Then the A.C. component of anode current is

$$i_a = \mu e_g/(R + \rho)$$

and the A.C. power W developed in the load is given by

$$\mu^2 R e_g^2/(R + \rho)^2$$

and, for a given value of e_g, will be a maximum when $R = \rho$. The ratio $\sqrt{W}/e_g$, which is independent of the input E.M.F. is termed the *Power Sensitivity* of the triode, and it is clear that the load resistance cannot be adjusted to give simultaneously both maximum power sensitivity ($R = \rho$) and maximum "undistorted" output ($R = 2\rho$). Fortunately the latter adjustment is not critical but, in any case, the primary consideration in the design of an output stage must be the selection of a suitable valve and the arrangement of conditions so that the required output may be obtained without undue distortion; the provision of high-power sensitivity is of secondary importance. When the type of valve and magnitude of load resistance have been fixed, there remains to be decided only the optimum value of grid bias. This will depend, of course, upon the available H.T. voltage, and is best determined by trial from graphical estimations of output and distortion for a few selected values. With valves designed for large outputs, it may happen that the steady anode current flowing when the optimum value

of grid bias is used, is so large that the power dissipated at the anode is greater than the valve can withstand. In this case, the steady grid potential must be made more negative, and it will then be advantageous to use a load resistance greater than twice the slope resistance of the valve.

Pentodes

We must now consider briefly the use in the output stage of valves other than triodes. In Chapter IV it was pointed out that, although the second grid in a screen-grid valve was originally introduced to reduce anode-grid capacity, it also produces a very beneficial change in the static characteristics of the valve. It seems probable, therefore, that it would be advantageous to use the screen grid even in valves where the reduction of anode-grid capacity is not very important. However, if this be attempted, a difficulty is encountered which may be explained as follows. In a triode, when the primary electron stream strikes the anode, it causes the latter to emit secondary electrons. Since the anode is always at a potential considerably higher than that of the grid, the electric field causes these electrons to return to the anode, and the result is the same as if they had never been emitted. When a screen-grid valve is used as a high-frequency amplifier the same thing occurs, since the potential of the anode does not vary by more than a few volts, and is always greater than that of the screen grid. On the other hand, if a screen-grid valve were used in other stages of a receiver, the variations of the potential of its anode would be much greater, and at certain instants the anode might be at a lower potential than the screen grid. Under these conditions, any secondary electrons emitted by the anode would be collected by the screen grid and the resulting decrease in anode current would cause serious distortion. The difficulty cannot be overcome by reducing the potential of the screen grid, since this would impair the efficiency

of the valve, but a solution has been found in the introduction of yet another grid between the screen grid and the anode. This additional grid may be connected either to the control grid or to the negative end of the filament, and, since in either case its potential will always be lower than that of the anode, it will prevent the escape of secondary electrons from the anode to any other electrode. A valve constructed on these lines is known as a *Pentode.* Although its anode current-grid voltage characteristics appear very similar to those of a triode of high amplification factor, its anode current–anode voltage

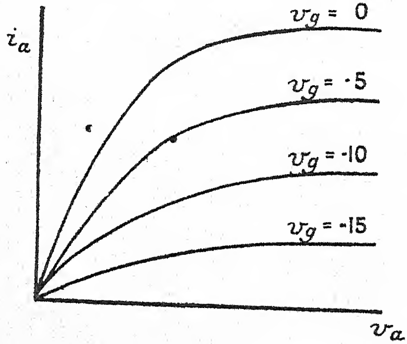

FIG. 41.

characteristics shown in Fig. 41, indicates that there is considerable difference between pentodes and triodes.

Hitherto pentodes have been chiefly used in output stages, and it is this use which we shall consider here. Unfortunately, the dynamic characteristic of a pentode with resistance load does not approximate to a parabola, so the geometrical analysis developed for a triode cannot be applied to the present case. It is therefore usual to investigate the capabilities of a pentode by a direct experimental method as follows. A known alternating E.M.F. of sinusoidal wave-form and of any convenient frequency is applied between control grid and filament

of the pentode and a known resistance load is placed in the anode circuit. The alternating E.M.F. developed across this resistance, owing to amplitude distortion, will not in general be sinusoidal, and is applied to the input terminals of a calibrated sharply-tuned amplifier, the output E.M.F. from which can be measured. The amplifier is tuned in turn to the fundamental and various harmonics of the original E.M.F., and thus the magnitude of the amplitude distortion may be determined. For further details of this method reference should be made to a paper by Pidgeon and McNally (33). It is clear that the method also enables the A.C. power delivered by the pentode to the resistance load to be calculated, so that, by varying the conditions of the experiment, it is possible to determine the optimum values of load and grid bias for a given pentode and given H.T. voltage. The optimum load is usually found to be about one-third of the slope resistance of the pentode, and is more critical than in the case of a triode. If a comparison be made between a pentode and a triode which are identical except for the extra electrodes in the former, it is found that, for a given H.T. voltage, there is little difference in the "undistorted" power output of which each is capable when working with its optimum load. On the other hand, the power sensitivity of the pentode is two or three times as great as that of the triode.

It will be appreciated that amplitude distortion may be produced by the detector and L.F. stages of a receiver as well as by the output stage. Also, the methods of investigation which have been described in this chapter are clearly of quite general application.

CHAPTER VIII

CONCLUSION

IN Chapter III it was pointed out that considerable difficulty arises when an attempt is made to provide a receiver with adequate selectivity whilst avoiding side-band cut-off. It cannot be said that this difficulty has yet been satisfactorily solved, but we may indicate the two lines of development which have been partially successful. In the first method the single-tuned circuit in the aerial-earth system is replaced by a band-pass filter consisting usually of two tuned circuits coupled together. By this means it is possible to obtain a resonance curve approximating very much more nearly to the ideal shown in Fig. 8 (*a*). For further details of this scheme, the reader should consult the articles mentioned in the bibliography (34). The difficulty with this arrangement is that the two circuits must be tuned simultaneously, and this is accomplished by mounting the two variable condensers on the same shaft. Then any variation with time in the calibration of either of the condensers will seriously affect the characteristics of the filter.

In the second method a single circuit is used, and the tuning is made sufficiently sharp to provide adequate selectivity. The resulting side-band cut-off is remedied by placing in the L.F. amplifier a filter circuit with a characteristic which accentuates the high frequencies with respect to the low. If properly designed, such a filter can be made to eliminate frequency distortion, but

a difficulty arises in that, since the side-band cut-off will vary as the wave-length to which the aerial is tuned varies, and since, moreover, it will be changed enormously by any alteration of reaction, it follows that complete correction cannot be obtained unless the filter also be variable.

From what has been said in the foregoing chapters, it will be clear that the general procedure in designing a wireless receiver will be as follows. First of all, the type of detector to be used must be settled, since upon this will depend the input E.M.F. which must be applied to the detector stage to ensure distortionless operation of this stage. Next the design of the aerial-earth system and the H.F. stage can be undertaken. The H.F. amplification must be sufficiently great to enable the correct E.M.F. to be passed on to the detector stage, even when weak signals are being received. Consequently, when the receiver is tuned to receive stronger signals, the detector will be overloaded unless means be provided to enable the H.F. amplification to be reduced. When a screen-grid valve is employed, this may be done by reducing the potential of the screen. Recently, special valves have been introduced in which the amplification factor may be varied by changing the grid bias (35).

The next portion of the receiver to be considered is the output stage. The design of this will, of course, depend upon the type of loud speaker to be used and the volume of sound required therefrom, but once it has been settled, the input E.M.F. which must be applied to it will be known. Finally, the L.F. amplifier can be designed, since both the output E.M.F. from the detector and the input E.M.F. to the power stage are known.

BIBLIOGRAPHY

GENERAL REFERENCES

Thermionic Vacuum Tubes, by E. V. Appleton. London (1932).
The Physical Principles of Wireless, by J. A. Ratcliffe. London (1929).
Wireless Principles and Practice, by L. S. Palmer. London (1928).
Principles of Radio Communication, by J. H. Morecroft. New York (1927).
Die Elektronenröhren und Ihre Technischen Anwendungen, by H. G. Möller (3rd edition). Braunschweig (1929).
A Critical Review of Literature on Amplifiers for Radio Reception, prepared by H. A. Thomas and F. M. Colebrook (Radio Research Special Report No. 9, Dept. of Scientific and Industrial Research). London (1930).

TEXT REFERENCES

[1] van der Pol, *Phil. Mag.*, vol. 7, p. 477 (1929).
[2] E. V. Appleton, *Thermionic Vacuum Tubes*, London (1932).
[3] Nichols, *Phys. Rev.*. N.S., vol. 13, p. 404 (1919).
[4] Miller, *Scientific Papers of the Bureau of Standards*, No. 351 (1919).
[5] Hartshorn, *Proc. Phys. Soc. Lond.*, vol. 39, p. 108 (1927).
[6] Colebrook, *Experimental Wireless*, vol. 6, p. 486 (1929).
[7] Smith-Rose and Colebrook, *Experimental Wireless*, vol. 2, p. 207 (1925).
[8] Eckersley, *Jour. I.E.E.*, vol. 60, p. 581 (1922).
[9] Loewe and von Ardenne, *Jahrb. draht. Telegr.*, vol. 27, p. 19 (1926).
[10] Drake and Browning, *Proc. Inst. Rad. Eng.*, vol. 13, p. 767 (1925).
[11] McLachlan, *Experimental Wireless*, vol. 4, p. 597 (1927).
[12] Beatty, *Experimental Wireless*, vol. 5, p. 3 (1928).
[13] Butterworth, *Experimental Wireless*, vol. 6, p. 293 (1929).
[14] Smith-Rose, *Proc. Phys. Soc. Lond.*, vol. 34, p. 127 (1922).
[15] Hull and Williams, *Phys. Rev.*, N.S., vol. 27, p. 432 (1926).
Hull, *Phys. Rev.*, N.S., vol. 27, p. 439 (1926).

[16] Colebrook, *Experimental Wireless*, vol. 2, pp. 330, 394, 459 (1925).
[17] Appleton and Taylor, *Proc. Inst. Rad. Eng.*, vol. 12, p. 277 (1924).
[18] Chaffee and Browning, *Proc. Inst. Rad. Eng.*, vol. 15, p. 113 (1927).
[19] Colebrook, *Experimental Wireless*, vol. 2, pp. 865, 946 (1925); and vol. 3, pp. 34, 90 (1926).
[20] Ballantine, *Proc. Inst. Rad. Eng.*, vol. 16, p. 593 (1928).
[21] Warren, *Experimental Wireless*, vol. 6, p. 425 (1929).
[22] Ballantine, *Proc. Inst. Rad. Eng.*, vol. 17, p. 1153 (1929).
[23] Turner, *Experimental Wireless*, vol. 7, p. 371 (1930).
[24] Terman and Morgan, *Proc. Inst. Rad. Eng.*, vol. 18, p. 2160 (1930).
[25] *Wireless World*, vol. 21, pp. 623, 653 (1927).
Vincer-Minter, *Wireless World*, vol. 23, p. 517 (1928).
[26] von Ardenne and Heinert, *Jahrb. drahtl. Telegr.*, vol. 26, p. 52 (1925).
[27] Colebrook, *Experimental Wireless*, vol. 4, p. 195 (1927)
[28] Anderson, *Proc. Inst. Rad. Eng.*, vol. 15, p. 195 (1927).
[29] Carpenter, British Patent, No. 325833 (1928).
[30] Sowerby, *Wireless World*, vol. 26, No. 22, p. 548 (1930).
[31] Brown, *Proc. Phys. Soc. Lond.*, vol. 36, p. 218 (1923-24).
[32] Hanna, Sutherlin and Upp, *Proc. Inst. Rad. Eng.*, vol. 16, p. 462 (1928).
[33] Pidgeon and McNally, *Proc. Inst. Rad. Eng.*, vol. 18, p. 266 (1930).
[34] Uehling, *Proc. Radio Club of America* (Nov. 1929).
Sowerby, *Wireless World*, vol. 26, p. 214 (1930).
Bligh, *Experimental Wireless*, vol. 9, p. 61 (1932).
[35] Ballantine, *Proc. Inst. Rad. Eng.*, vol. 18, p. 2102 (1930).

ADDITIONAL REFERENCES

Articles on Superheterodyne Receivers

Armstrong, *Proc. Inst. Rad. Eng.*, vol. 12, p. 539 (1924).
Turner, *Experimental Wireless*, vol. 4, pp. 286, 339, 402, 568 (1927).
Sowerby, *Wireless World*, vol. 27, pp. 689, 712 (1930).

Articles on the Apparent Demodulation of a Weak Signal by a Strong One

Beatty, *Experimental Wireless*, vol. 5, p. 300 (1928).
Butterworth, *Experimental Wireless*, vol. 6, p. 619 (1929).
Appleton and Boohariwalla, *Wireless Engineer*, vol. 9, p. 136 (1932).

INDEX

PRINTED BY
JARROLD AND SONS LTD.
NORWICH